



**Etudes structurales par RMN des profils Saccharidiques
d'Héparanes sulfates et de leur régulation cellulaire :
Mise en place d'un protocole de marquage, de
purification et d'analyse de chaines entières**

Mathieu Pegeot

► **To cite this version:**

Mathieu Pegeot. Etudes structurales par RMN des profils Saccharidiques d'Héparanes sulfates et de leur régulation cellulaire : Mise en place d'un protocole de marquage, de purification et d'analyse de chaines entières. Biologie structurale [q-bio.BM]. Université de Grenoble, 2014. Français. NNT : 2014GRENV056 . tel-01315172

HAL Id: tel-01315172

<https://theses.hal.science/tel-01315172>

Submitted on 12 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 7 août 2006

Présentée par

Mathieu PEGEOT

Thèse dirigée par **Pierre GANS** et **Rabia SADIR**

préparée au sein des **Laboratoires de RMN Biomoléculaire et Structure et Activité des Glycosaminoglycanes** à l'Institut de **Biologie Structurale – UMR 5075**
dans l'**École Doctorale Chimie et Sciences du Vivant**

Etudes Structurales par RMN des Profils Saccharidiques d'Héparanes Sulfates et de leur Régulation Cellulaire : Mise en Place d'un Protocole de Marquage, de Purification et d'Analyse de Chaines Entières

Thèse soutenue publiquement le **11 Décembre 2014**,
devant le jury composé de :

Pr Franck FIESCHI

Professeur à l'Université Grenoble Alpes, Président

Dr Isabelle LANDRIEU

Chargée de Recherche CNRS à l'Université de Lille 1, Rapporteur

Pr Dulce PAPY-GARCIA

Professeur à l'Université Paris Est, Rapporteur

Pr Sylvie RICARD-BLUM

Professeur à l'Université Claude Bernard Lyon 1, Examineur

Dr Pierre GANS

Institut de Biologie Structurale, directeur de thèse, Examineur

Dr Rabia SADIR

Institut de Biologie Structurale, co-encadrante de thèse, Examineur



Remerciements

En premier lieu je tiens à remercier les membres du jury, Isabelle Landrieu et Dulce Papy-Garcia pour avoir accepté d'être rapporteurs de mon travail de thèse ainsi que Sylvie Ricard-Blum, Franck Fieschi, Rabia Sadir et Pierre Gans en qualité d'examinateurs.

Rabia, Pierre, l'aventure à trois a commencé dès le M2 il y a bientôt 4 ans où vous m'avez ouvert les portes du laboratoire, de la RMN et des GAG. J'ai pris plaisir à évoluer à vos côtés tout au long de ces années, dans votre bonne humeur et bienveillance. Merci d'avoir été là dans cette épreuve qu'est la thèse et de vous être rendus disponibles à chaque fois que j'en avais besoin, je ne l'oublie pas, tout comme je n'oublie pas l'énergie que vous avez mis pour corriger mon manuscrit.

Hugues, Jean-Pierre, je vous remercie profondément pour l'investissement dont vous avez fait preuve pour compléter ma formation et d'avoir suivi le projet de près au travers des différentes réunions et aux CST auxquels vous avez toujours assisté sans oublier la rédaction de l'article. Votre gentillesse et votre optimisme m'ont toujours permis d'aller de l'avant, même lorsque j'avais l'impression de faire du surplace.

Cédric, cette thèse aussi c'est à toi que je la dois. Depuis le premier jour tu as su être d'une grande aide grâce à ta palette de biochimiste très fournie. J'ai appris beaucoup à tes côtés avec tes nombreux conseils, j'ai également passé de très bons moments en ta compagnie au laboratoire où tu faisais rimer travail, entrain et bon vivre. De la colonne de chromatographie à l'utilisation de Linux en passant par tout le reste tu as été là pour m'aider et tu t'es toujours rendu disponible. Pour tout cela Cédric, je t'en remercie sincèrement et te souhaite le meilleur.

Je remercie également Sabine Hediger et Hiroki Takahashi qui m'ont accueilli au sein de leur laboratoire pour quelques expériences avec une grande sympathie. J'ai apprécié être dans un autre cadre que mon propre laboratoire.

Merci également à tout le SAGAG où il a fait bon vivre : Romain avec toujours les bonnes *vibes* et la bonne humeur, toujours très musicale ! Amal, je te souhaite une bonne fin de thèse avec un bel avenir pour toi. Damien, même si ça a souvent été l'embouteillage au niveau de l'AKTÄ tu nous manquais 3 semaines sur 6, merci pour ta sympathie et je te souhaite une suite également heureuse tant au niveau de tes concentrations en protéines/sucres que dans ta vie perso. Enfin Yoan, je te souhaite également le meilleur pour l'avenir comme pour le présent, et Séb qui est déjà parti. Je n'oublie pas non plus Bridgette et Els, vous étiez deux beaux exemples à suivre tant sur le plan scientifique que sur le plan humain. Votre dynamisme a été moteur pour moi, et les sorties en dehors toujours des moments de joie. En tant que mères, épouses et femmes de science, je vous souhaite d'être comblées.

Merci aussi au LRMN, Adrien, Bernhard, Dominique, Jean-Pierre, Paul pour le travail impressionnant de qualité que vous faites sur les spectromètres pour la communauté et pour le temps que vous passez aussi à nous former sur le plan théorique. Catherine tu as le cœur sur la main et tu te soucies de tous tes étudiants dont j'ai fait partie, je te remercie de cette bienveillance et de ta profonde générosité, sache que j'en ai été touché. Nicolas, au-delà de tes compétences d'attribueur fou et de gamer invétéré tu as toute mon admiration pour

l'énergie que tu mets dans ton travail, je te souhaite une bonne fin de thèse ainsi qu'un parcours professionnel et personnel qui te comblera. Enfin je souhaite bonne route à ceux qui restent et qui évolueront sous d'autres horizons avec qui j'ai pu passer ou aurais pu passer sans l'ombre d'un doute de bons moments : Lionel, Elodie, Rime, Pavel, Peixiang, Hugo, Guillaume, Katharina, Caroline, Karine, Vilius, Jia-Ying... Et Lauriane, un grand merci pour tout ton pep's communicatif qui a été fédérateur, je te souhaite une belle suite dans la Belle Province et n'importe où tu seras.

Et que dire de l'IRPAS, mon LEM voisin avec qui j'ai plaisir à discuter ou à n'échanger qu'un sourire. Isabelle, Véronique, Pascale, Evelyne et enfin Nicole je vous remercie d'être si gentilles et attentionnées. Vous êtes des personnes en or qui avaient été là aux bons moments pour moi. Je vous témoigne ici toute mon affection et ma sympathie. Bien évidemment je ne t'oublie pas Monique, je te souhaite sincèrement la plus belle des retraites en recevant tout l'amour que tu as toujours donné sans limite. Tu (me) manques beaucoup. Mélanie aussi je pense à toi et te remercie de m'avoir donné un bel exemple, celui d'une thésarde avec un caractère fort et une gentillesse débordante. Que le ciel de ton futur soit clair et sans nuage, c'est-à-dire presque tout le contraire de ton ciel écossais ;-)

Mes pensées vont également à un trio de l'équipe DIR de chic et de choc : Dominique, Linda et Blandine. Je vous regretterai énormément tant par la bonne humeur et la gentillesse dont vous faites preuve au quotidien mais également par votre efficacité dans l'administratif. Pour tout cela vous avez toute mon admiration. Alain aussi tu vas me manquer avec toutes tes plaisanteries, et même si je connais presque aussi bien que toi tes collections tant tu aimes à en parler, j'espère que tu trouveras toujours quelqu'un avec qui partager cela. Enfin, Fabrice, Sylvie, Carole je vous souhaite également le meilleur.

Je tiens également à remercier les amis qui étaient là, qui sont partis, mais que je n'ai jamais quittés. Aurélie, tu es une personne formidable et je te souhaite un destin à ta hauteur. Ma Josiane, si mauviiiise et si drôle ! Que de bons souvenirs, quelle aventure ! J'ai quand même un peu l'impression de rêver ! Hé hé hé hé Sarah, le concentré de vitalité ! On se sera bien marré au SAGAG. Jens, Mathilde, les pauses *papples* et *bananas* de 16h me manquent depuis que vous êtes partis. Tout le bonheur pour vous pour la suite, et vous serez toujours les bienvenus n'importe où vous serez, et veel geluk Mathilde !

Maintenant passons à la Team Porquerolles (élargie), avec de grandes et belles majuscules. Au-delà du fait que nous avons été dans le même bateau au sens propre comme au sens figuré, notre condition de thésards nous a rapproché, mais pas que. En effet des liens étroits indéfectibles se sont tissés. Grâce à ces liens, un véritable tissu amical s'est créé plus largement, et vous avez tous été là pour moi dans tous les moments de joie comme dans les épreuves. Votre amitié est gravée en moi au fer rouge, ne plus vous avoir au quotidien me rend déjà nostalgique. Big up à tous donc : Louise, Widade, H2K, Hicham, Yann, Jules, Gianluca, Didier, Zsófia, Johannes, Vanessa, Marie, Coraline, Pierre, Michel, Marlène, Vitaliy, Ivan, Lorraine, Mylène... Et la team repas avec en plus Delphine, Céline et Florine. Une grosse pensée aussi à toi Elodie, une femme forte qui sait où elle va. Ton tempérament de leader te mènera là où tu le souhaites et te remplira de bonheur, j'en suis convaincu. Ziad mon camarade, mon collègue, mon frère du Liban, mon ami, je te souhaite toute la réussite pour

le futur. Audrey bonne continuation à toi également, tes qualités précieuses te mèneront plus loin sur le chemin du bonheur même si tu y marches déjà.

Il y a aussi la fine équipe de la fac que je tiens à remercier pour les (très) bons moments que j'ai passé avec vous, et ce malgré le fait que chacun ait été dans ses pipettes : Alice, Clément, Alexandre, Jessica, Vincent vous êtes une soupape formidable !

Pour survivre à une thèse il faut aussi de bons amis, de très bons amis, en fait, les meilleurs et je vous en remercie : ma roommie qui m'a supporté dans toutes mes bêtises et dans tous les autres moments de vie : Amélie ainsi que Tristan ; Priscille mon pilier, mon oxygène ; Mathieu mon Dupont qui est là depuis tellement d'année ; Jojo, ami de la première heure aussi et Jach; Juju-chan (la reine de la fête à Paris comme à Nice) ; Elsa et Séb (toujours présents, amusants et réconfortants) ; Julie et Séb (force et honneur à vous aussi) ; Lise, tu me scieras toujours en deux tant tu es drôle ; Caro et Jimmy les superhéros du quotidien ; Jérémie mon ami Limougeaud ; Clément ; Anthony.

Et puis il y a la famille, à qui cette thèse est dédiée. C'est un merci viscéral, vous m'avez toujours poussé à donner le meilleur de moi-même, m'avez toujours soutenu dans mes différentes entreprises, avaient été là depuis le début. Si j'en suis là c'est grâce à vous, et je vous témoigne tout mon amour : Moscou, Emilie, Thomas, Laura, mes parents et ceux qui me comblent et qui assureront la relève Evan et Jordan. Avec vous, rien ne saurait me manquer.

Table des matières

Remerciements	1
Table des matières	5
Abréviations	11
Introduction	15
A. La Matrice extracellulaire (MEC)	15
1. Généralités.....	15
2. Les collagènes	16
3. Les glycoprotéines matricielles	17
4. Les molécules élastiques	18
5. Les protéoglycanes (PG)	19
B. Les Protéoglycanes.....	20
1. Historique	20
a) Découverte du premier GAG : histoire autour du « chondroitinschwefelsäure »	20
b) Découverte des protéines cœur du PG	21
2. Les protéines cœur	21
a) Classification	21
(1) Les PG de la matrice extracellulaire.....	22
(a) Les SLRP (<i>Small Leucine Rich Proteoglycans</i>)	22
(b) Les hyalectans (ou lecticans)	23
(c) Les autres PG de la MEC	24
(2) Les PG membranaires.....	24
(a) Le CD44.....	25
(b) Le bétaglycan	26
(c) Les syndécans	26
(d) Les glypicans	28
(3) Les PG intracellulaires.....	29
(4) Les PG circulants.....	31
b) Synthèse et tri	31
c) Le catabolisme des PG	33
C. Les GAG et leur métabolisme	35
1. Généralités.....	35
2. L'acide hyaluronique (HA)	36
a) Synthèse	36
b) Structure	38
c) Catabolisme	38
3. Les kératanes sulfates (KS)	39

a)	Synthèse	39
(1)	Initiation.....	39
(2)	Elongation.....	40
(3)	Terminaison	40
b)	Structures et modifications.....	41
4.	Les chondroïtines sulfates (CS), héparanes sulfates (HS) et Héparine (Hp).....	42
a)	Le tétrasaccharide linker	42
b)	Les chondroïtines sulfates (CS).....	43
(1)	Biosynthèse.....	43
(2)	Structures et modifications	44
c)	Les héparanes sulfates (HS) et l'héparine (Hp)	46
(1)	Biosynthèse.....	47
(2)	Structures et modifications	48
5.	Le catabolisme des GAG <i>in vivo</i>	52
D.	Rôles des PG	54
1.	Rôles de la protéine cœur... ..	54
a)	... dans le choix de la chaîne de GAG	54
b)	... dans des phénomènes de signalisation cellulaire	55
c)	... dans l'architecture matricielle	55
2.	Rôles physiologiques des GAG cellulaires	55
a)	Les HS/Hp peuvent provoquer des changements conformationnels : cas de l'héparine dans la coagulation.....	56
b)	Les HS/CS dans la rétention et protection de protéines	58
c)	Les HS comme corécepteurs	58
d)	Les HS dans la formation de gradients de chimiokines et dans l'inflammation ..	59
3.	Les GAG dans des phénomènes pathologiques	61
a)	GAG et infections.....	61
b)	GAG et cancer	61
c)	Les maladies génétiques.....	63
E.	Intérêts de l'étude des GAG	63
1.	Les GAG comme médicaments.....	64
a)	Dans l'hémostase.....	64
b)	Contre l'arthrose.....	64
c)	Comme anticancéreux	64
2.	Industrie des GAG et PG dans la Santé.....	66
a)	L'industrie de l'héparine	66
b)	Le scandale de l'OSCS (oversulfated chondroitin sulfate)	67
c)	L'industrie de l'acide hyaluronique	68

F. Etudes structurales des GAG.....	69
1. Détermination de la structure primaire des GAG.....	70
a) La dépolymérisation des HS	70
(1) La dépolymérisation chimique	70
(2) La dépolymérisation enzymatique.....	70
b) Etudes par électrophorèse.....	71
(1) Etudes par électrophorèse capillaire (CE)	71
(2) Etudes par FACE (<i>Fluorescence-Assisted Carbohydrate Electrophoresis</i>)...	72
c) Etudes par chromatographie	72
d) Etudes par spectrométrie de masse (MS)	74
2. Analyses dynamiques et tridimensionnelles des GAG.....	75
a) Etudes par microscopie	75
b) Etudes par diffraction aux rayons X.....	76
c) Etudes par diffusion des rayons X aux petits angles (SAXS)	78
d) Etudes de la dynamique moléculaire (MD).....	79
3. La RMN, une technique de choix.....	80
a) Principes de la RMN	80
b) Etudes de la structure et de la dynamique des GAG par RMN en solution	86
c) Enrichissement isotopique des GAG.....	89
d) Etudes d'interaction HS-protéine par RMN	91
e) Etudes des GAG par RMN en phase solide	92
Problématique et objectif du travail	94
Matériel et Méthodes.....	97
A. Culture cellulaire	97
1. Cellules HeLa adhérentes	97
2. Cellules HeLa en solution et taux de mortalité.....	97
3. Fibroblastes de souris embryonnaires (MEF)	97
4. Cellules Caco-2	98
5. Marquage au ¹³ C et amplification des cellules	98
a) Marquage.....	98
b) Flasques de culture	98
B. Cytométrie en flux.....	99
1. Préparation des cellules	99
2. Analyse.....	99
C. Digestions enzymatiques.....	100
1. ...des surfaces cellulaires aux protéases	100
a) Trypsine.....	100
b) Papaïne	100

c) Pronase	101
2. ...des extraits cellulaires aux nucléases	101
a) Benzonase.....	101
b) DNase I.....	101
D. Purification par chromatographie échangeuse d'anions.....	102
1. Chargement des échantillons.....	102
a) Après digestion aux protéases	102
b) Après digestion aux nucléases.....	102
2. Elution en <i>step</i> de sel.....	102
3. Elution en gradient de sel	103
E. Préparation des extraits d'HS	103
1. Dessalage.....	103
a) Système PD-10.....	103
b) Système Hitrap Sephadex.....	103
c) Dialyse.....	104
2. Lyophilisation.....	104
3. Préparation des HS avec réduction d'étapes pré-analytiques.....	104
F. Dosage des HS	104
1. Dosage par carbazole	104
2. Dosage par Blyscan TM	105
3. Dosage par RMN.....	105
a) Analyse de l'échantillon de référence	105
b) Analyse de l'extrait d'HS	106
G. Analyse par RMN liquide des extraits d'HS	106
1. Expériences de corrélation	106
2. Quantification de la modification des HS	107
H. Etudes des cellules entières	107
1. Préparation des cellules	107
a) Cellules entières	107
b) Cellules dépourvues d'HS, CS et N-glycans.....	107
2. Analyse par RMN solide	107
a) Par rotation à l'angle magic MAS (Magic Angle Spinning).....	107
b) Par polarisation dynamique nucléaire (DNP).....	108
I. Etudes des membranes plasmiques isolées	108
1. Préparation des membranes.....	108
2. Etudes par ssNMR.....	108
3. Solubilisation des membranes	109
a) Triton X-100.....	109

b) OctylPOE	109
c) DPC	110
4. Digestion des HS et CS	110
5. Analyse par RMN liquide	110
a) HSQC	110
b) HCCH-COSY (COrrrelation SpectroscopY).....	110
c) HCCH-TOCSY (TOtal Correlation SpectroscopY).....	110
Résultats et Discussion.....	113
A. Mise en place d'un protocole de marquage et de purification d'HS entiers cellulaires pour leurs analyses RMN	113
1. Vérification de l'expression en HS et CS des cellules HeLa par cytométrie en flux	114
2. Conditions de culture et de marquage cellulaire	115
a) Conditions de culture	115
b) Détermination des conditions de marquage	116
3. Extraction, purification et analyse RMN des HS	117
a) Digestion des surfaces cellulaires	117
(1) Digestion des surfaces cellulaires à la trypsine	117
(2) Digestion des surfaces cellulaires à la papaine.....	118
(3) Digestion des surfaces cellulaires à la pronase.....	118
b) Purification et analyse RMN des digestats cellulaires	119
(1) Purification et analyse RMN de la digestion à la trypsine des surfaces cellulaires	119
(2) Purification et analyse RMN après digestion à la papaine des surfaces cellulaires	120
(3) Purification et analyse RMN après digestion au cocktail pronase des surfaces cellulaires	122
c) Digestion des principaux contaminants, les acides nucléiques	123
(1) Digestion à la benzonase, purification et analyse RMN d'HS cellulaires	124
(2) Digestion à la DNase I, purification et analyse RMN d'HS cellulaires	125
4. Amélioration des conditions d'extraction et de purification des HS cellulaires.....	127
a) Choix des flasques de culture	127
(1) Système HYPERflask™ (Corning Life Science®).....	127
(2) Système T600 (EDM Millipore).....	128
b) Optimisation de la méthode d'élution des HS par chromatographie échangeuse d'anions.....	130
(1) Éluion en <i>step</i>	130
(2) Éluion en gradient de sel	131
c) Réduction d'étapes pré-analytiques	133
5. Quantification des HS	134

a)	Dosage des acides uroniques par la méthode au carbazole	134
b)	Dosage des GAG sulfatés par la méthode Blyscan TM commercial	135
c)	Dosage des HS par RMN	136
B.	Analyses par RMN en solution des HS purifiés issus des cellules HeLa	138
1.	Influence de la température d'analyse	138
2.	Attribution des signaux	139
3.	Quantification des signaux	141
a)	La N-sulfatation.....	141
b)	L'épimérisation des acides glucuroniques	142
c)	La 2-O-sulfatation des acides iduroniques	143
d)	La 6-O-sulfatation des glucosamines	143
e)	La longueur des chaînes	144
f)	2-O-sulfatation des GlcA, 3-O-sulfatation des GlcN et glucosamines non-substituées	144
g)	Analyse des unités disaccharidiques	145
C.	Etudes de polysaccharides d'HS issus de deux types cellulaires : application du protocole et analyse par RMN liquide.....	146
1.	Les fibroblastes de souris embryonnaires (MEF)	146
a)	Détermination par cytométrie en flux de la présence d'HS au niveau des surfaces cellulaires	146
b)	Marquage et purification	147
c)	Analyse RMN de l'évolution du profil saccharidique.....	148
2.	Les cellules Caco-2	151
a)	Détermination par cytométrie en flux de la présence d'HS au niveau des surfaces cellulaires	151
b)	Marquage et purification	153
c)	Analyse RMN de l'évolution du profil saccharidique.....	153
D.	Etude des HS dans le contexte du protéoglycane.....	156
1.	Etudes des sucres à partir de cellules entières	156
2.	Premières analyses par polarisation nucléaire dynamique (DNP)	159
3.	Etudes des sucres dans les membranes plasmiques isolées de cellules HeLa.....	160
4.	Etudes par RMN liquide des membranes de cellules HeLa solubilisées dans des détergents	163
	Discussion générale, Conclusion et Perspectives.....	169
	Annexes	177
	182
	183
	184
	Bibliographie	213

Abréviations

ADN(c)	Acide DésoxyriboNucléique (complémentaire)
AFM	Atomic Force Microscopy
ARN	Acide RiboNucléique
ATIII	Anti-Thrombine III
AU	Absorbance Unit
BMP-4	Bone Morphogenetic Protein 4
BSA	Bovine Serum Albumin
CE	Capillary Electrophoresis
CHO	Chinese Hamster Ovary
ChPF	Chondroïtin Polymerization Factor
ChSy	Chondroïtine Synthétase
CP	Cross-Polarization
CS	Chondroïtine Sulfate
Da/kDa	Dalton/kilodalton
DEAE	DiEthylAminoEthyl
DMEM	Dulbecco's Modified Eagle Medium
DNP	Dynamic Nuclear Polarization
DO	Densité Optique
dp	degré de polymérisation
DPC	n-DodécylPhosphoCholine
DS	Dermatane Sulfate
EDTA	Ethylène Diamine Tétra Acétique
EGFR	Epidermal Growth Factor Receptor
ERGIC	Endoplasmic Reticulum – Golgi Intermediate Compartment
ERT	Enzyme Replacement Therapy
EXT	Exostosine
EXTL	Exostosine-Like
FACE	Fluorescence-Assisted Carbohydrate Electrophoresis
FACIT	Fibril-Associated Collagen with Interrupted Triple helix
FACS	Fluorescence-Activated Cell-Sorting
FGF	Fibroblast Growth Factor
FGFR	Fibroblast Growth Factor Receptor
FID	Free Induction Decay
FITC	Fluorescein IsoThioCyanate
FPLC	Fast Protein Liquid Chromatography
FSC	Forward-Scattered light
FT	Fourier Transform
Fuc	Fucose
GAG	GlycosAminoGlycane
Gal	Galactose
GalNAc(4S/6S)	<i>N</i> -Acétyle Galactosamine (4- <i>O</i> -sulfate/6- <i>O</i> -sulfate)
GalNAcT	<i>N</i> -AcétylGalactosaminylTransférase
GalT	GalactosylTransférase

GDNF	Glial cell line-Derived Neurotrophic Factor
Glc	Glucose
GlcA(2S)	Acide Glucuronique (2- <i>O</i> -sulfate)
GlcAT	GlucuronylTransférase
GlcNAc(6S)	<i>N</i> -Acétyl Glucosamine (6- <i>O</i> -sulfate)
GlcNAcT	<i>N</i> -AcétylGlucosaminylTransférase
GlcNS(6S)	<i>N</i> -sulfate Glucosamine (6- <i>O</i> -sulfate)
Gpc	Glypican
GPI	GlycosylPhosphatidylinositol
HA	Hyaluronic Acid
HAS	Hyaluronic Acid Synthase
HBD	Heparin Binding Domain
HCII	Heparin Cofactor II
Hep	Héparinase
HexA	Hexuronic Acid
HMWHA	High Molecular Weight Hyaluronic Acid
Hp	Héparine
HPLC	High Performance Liquid Chromatography
HS	Héparane Sulfate
HSQC	Heteronuclear Single Quantum Coherence
HSV	Herpes Simplex Virus
HSxST	Heparan Sulfate x- <i>O</i> -SulfoTransférase
IdoA(2S)	Acide Iduronique (2- <i>O</i> -sulfate)
IFN-γ	Interféron gamma
IgM	Immunoglobuline de type M
INEPT	Insensitive Nuclei Enhanced by Polarization Transfer
KO	Knock-Out
KS	Kératane Sulfate
LMWH	Low Molecular Weight Heparin
LMWHA	Low Molecular Weight Hyaluronic Acid
LRR	Leucine Rich Repeat
mAb	monoclonal Antibody
Man	Mannose
MAS	Magic Angle Spinning
MD	Molecular Dynamics
MEC	Matrice ExtraCellulaire
MEF	Mouse Embryonic Fibroblast
MPS	MucoPolySaccharidose
MS	Mass Spectrometry
NDST	<i>N</i> -Désacétylase <i>N</i> -SulfoTransférase
NEAA	Non-Essential Amino Acids
NeuAc	Neuraminic Acid
nOe	nuclear Overhauser effect
NOESY	Nuclear Overhauser Effect Spectroscopy
octylPOE	n-octylPolyOxyEthylène
OSCS	OverSulfated Chondroitin Sulfate
PAPS	3'-PhosphoAdénosine-5'-PhosphoSulfate

PBS	Phosphate Buffered Saline
PDGF	Platelet-Derived Growth Factor
PFA	ParaFormAldéhyde
PG	ProtéoGlycane
ppm	partie par millions
RE	Réticulum Endoplasmique
RF	RadioFréquence
RMN	Résonance Magnétique Nucléaire
ROS	Reactive Oxygen Species
RT	Room Temperature
SA	Syalic Acid
SAX	Strong Anion Exchange
SAXS	Small-Angle X-ray Scattering
Sdc	Syndécan
SDF-1 α	Stromal cell-Derived Factor 1 alpha
sGAG	sulfoGlycosAminoGlycanes
Shh	Sonic hedgehog
SLRP	Small Leucine-Rich Proteoglycan
SSC	Side-SCattered light
ssNMR	solid state Nuclear Magnetic Resonance
Sulf	Sulfatase
SVF	Sérum de Veau Fœtal
TGF- β	Transforming Growth Factor beta
TGN	TransGolgi Network
TLR	Toll-Like Receptor
TNF- α	Tumor Necrosis Factor alpha
TOCSY	TOTAL Correlation Spectroscopy
UDP	Uridine DiPhosphate
UFH	UnFractionated Heparin
UST	Uronyl-2- <i>O</i> -SulfoTransférase
VEGF	Vascular Endothelial Growth Factor
VIH	Virus de l'Immunodéficience Humaine
Wnt	Wingless type
Xyl	Xylose
XylT	XylosylTransférase

Introduction

A. La Matrice extracellulaire (MEC)

1. Généralités

La matrice extracellulaire (MEC) est l'ensemble des macromolécules à l'extérieur des cellules d'un organisme sur laquelle elles peuvent se fixer, de manière transitoire ou durable. La MEC forme un réseau d'ancrage également pour les protéines, sucres, lipides et acides nucléiques qui jouent des rôles essentiels dans les phénomènes cellulaires physiologiques et/ou pathologiques (Figure 1). Cette matrice, en plus d'une fonction cargo en molécules d'intérêts, est un support mécanique qui permet de faire le lien entre les interactions intercellulaires et cellule-environnement afin de s'adapter efficacement aux différents besoins des cellules de l'organisme ainsi qu'aux évolutions de son environnement. Ce support physique est essentiel dans les phénomènes de migration cellulaire et notamment lors de l'embryogenèse.

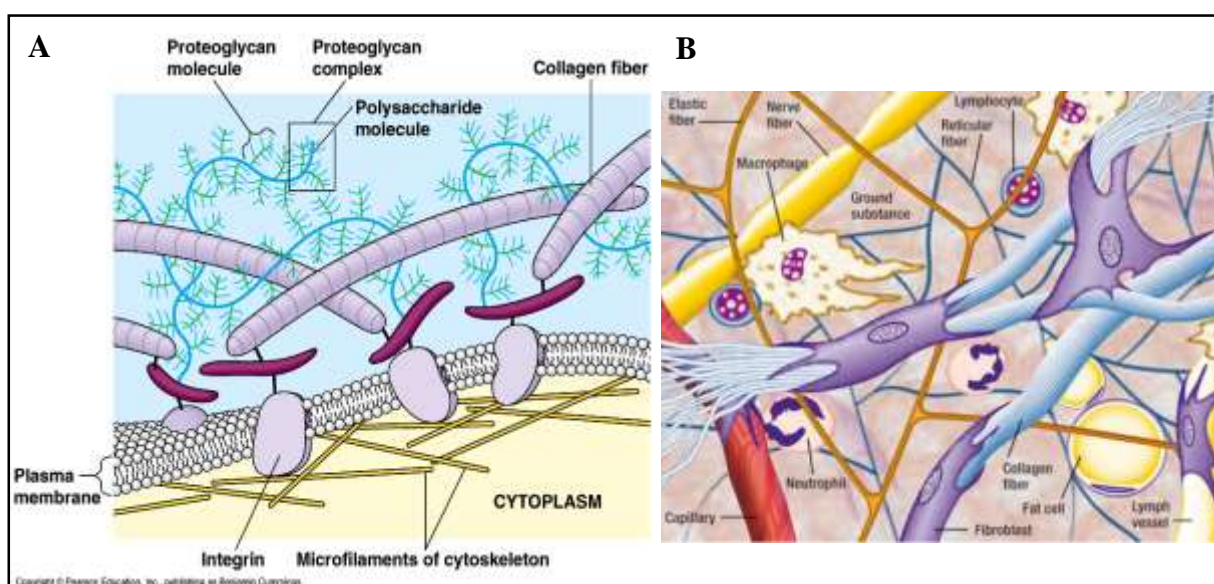


Figure 1 : Schémas représentant la matrice extracellulaire et leurs composantes vues au niveau des surfaces cellulaires (A) et au sein d'un tissu (B). D'après Pearson Education Inc.

L'apparition de la matrice extracellulaire au cours de l'Évolution est liée à l'apparition des premières espèces animales pluricellulaires (éponges) il y a plus 500 millions d'années lors de l'Édiacarien, dernière période géologique du Pré-Cambrien (Gehling & Rigby 1996). Cette matrice, aujourd'hui répandue chez tous les métazoaires, a été la clé dans l'arrangement des cellules en tissus et à l'élaboration de nouvelles structures. Au cours de l'évolution, les quantités relatives des molécules qui la composent ont varié tout comme son organisation, aboutissant à une diversité de formes (par exemple chez les vertébrés avec la calcification de certaines MEC) et de spécificités.

La composition de la MEC regroupe 4 grandes « familles » de molécules (collagènes, glycoprotéines, molécules élastiques et protéoglycanes) qui permettent au tissu d'assurer ses

rôles de soutien mécanique, de rétention/distribution en protéines d'intérêt pour les cellules environnantes, de lubrification des tissus et cohésion cellulaire (Figure 1).

2. Les collagènes

Les collagènes forment une famille de protéines comprenant 28 membres recensés à ce jour et représentent plus d'un quart de la masse protéique totale chez l'Homme. Ces protéines sont sécrétées principalement par les fibroblastes et possèdent une organisation structurale typique qui donne aux collagènes leur propriété de résistance à la traction et à la tension. En effet, les collagènes partagent un motif tripeptidique caractéristique de séquence Gly-X-Y où Gly est une glycine, X une proline et Y une 4-hydroxyproline le plus souvent. Cette séquence forme un domaine en triple hélice, c'est-à-dire qu'en s'assemblant avec 2 autres partenaires du même type (homotrimère) ou de la même famille (hétérotrimère) les 3 protéines vont former une triple hélice. Cette structure superenroulée positionne les chaînes latérales des acides aminés vers l'extérieur de l'hélice, ce qui va permettre l'interaction avec d'autres molécules. Les triples hélices peuvent interagir avec d'autres triples hélices de collagène pour former des fibres de collagène. Leur composition varie en fonction du tissu dans lequel elles sont localisées. Par exemple, les collagènes II, III, IX et XI sont retrouvés associés dans le cartilage (Wu et al. 2010) alors que l'association des collagènes I et V est retrouvée dans la cornée (Bruckner 2010). La longueur des fibres varie selon le tissu et le type de collagène engagé dans la triple hélice et sera, par exemple, plus importante dans un tendon qu'au niveau de la peau. Mais il existe également d'autres modes d'association des collagènes, comme les assemblages en réseau, en filament, et divers arrangements de type fibre (fibre d'ancrage, simple ou avec interruption de la triple hélice) qui dépendent du type de collagène (Figure 2) (Ricard-Blum 2011). Les collagènes peuvent également être porteurs de chaînes de

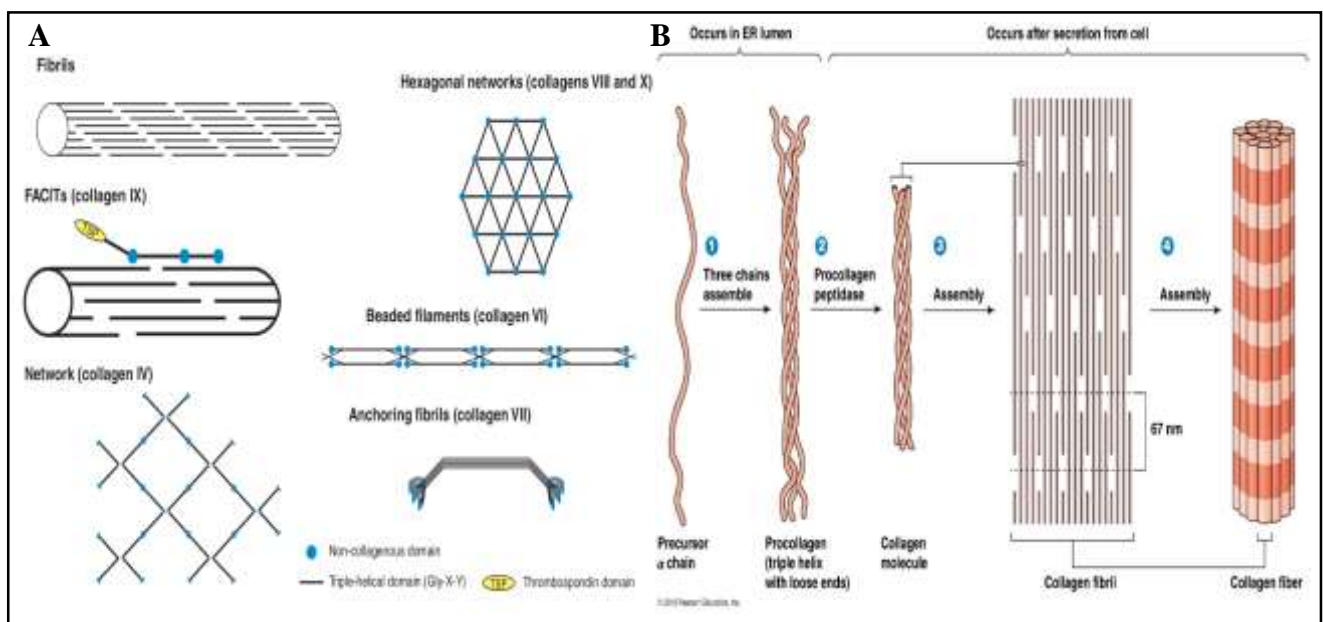


Figure 2 : Différents assemblages (A) et mode d'assemblage supramoléculaires (B) de collagènes. D'après Ricard-Blum 2011 et Pearson Education Inc.

glycosaminoglycanes (GAG) et donc former un protéoglycane, c'est le cas par exemple du collagène XVIII. La coupure de ce dernier permet de relarguer un fragment soluble appelé endostatine présentant une activité anti-angiogénique (Seppinen & Pihlajaniemi 2011).

La formation de ces superstructures insolubles est encore mal connue et ne dépend pas que de la proportion en différents collagènes pour aboutir à la structure voulue *in vitro* (Bruckner 2010). En effet, les associations des triples hélices de collagènes sont initiées, guidées ou influencées par d'autres collagènes (Wenstrup et al. 2004), par différents composants de la MEC comme les protéoglycanes (Chakravarti et al. 2000) (Kalamajski & Oldberg 2010), la fibronectine ou par des protéines transmembranaires comme les intégrines par exemple (Kadler et al. 2008).

3. Les glycoprotéines matricielles

Plus de 200 glycoprotéines sont connues et forment une famille comportant des protéines de haut poids moléculaire (jusqu'à 1000kDa) auxquelles sont attachés de manière covalente des oligosaccharides, d'où le terme de « glycoprotéine ». Elles peuvent être ubiquitaires (cas de la fibronectine) ou tissu-spécifiques (cas des laminines retrouvées dans les membranes basales). Leur rôle est essentiel dans l'adhésion avec les molécules de la MEC et de la surface cellulaire comparé au rôle plutôt structural des collagènes.

Les glycoprotéines matricielles ont des structures (ou modules) connues favorisant la multimérisation et les interactions avec les autres protéines exposées dans la MEC. La fibronectine, par exemple, forme un dimère qui possède des motifs d'interaction avec les collagènes de la matrice comme avec les intégrines cellulaires (Figure 3). L'interaction avec

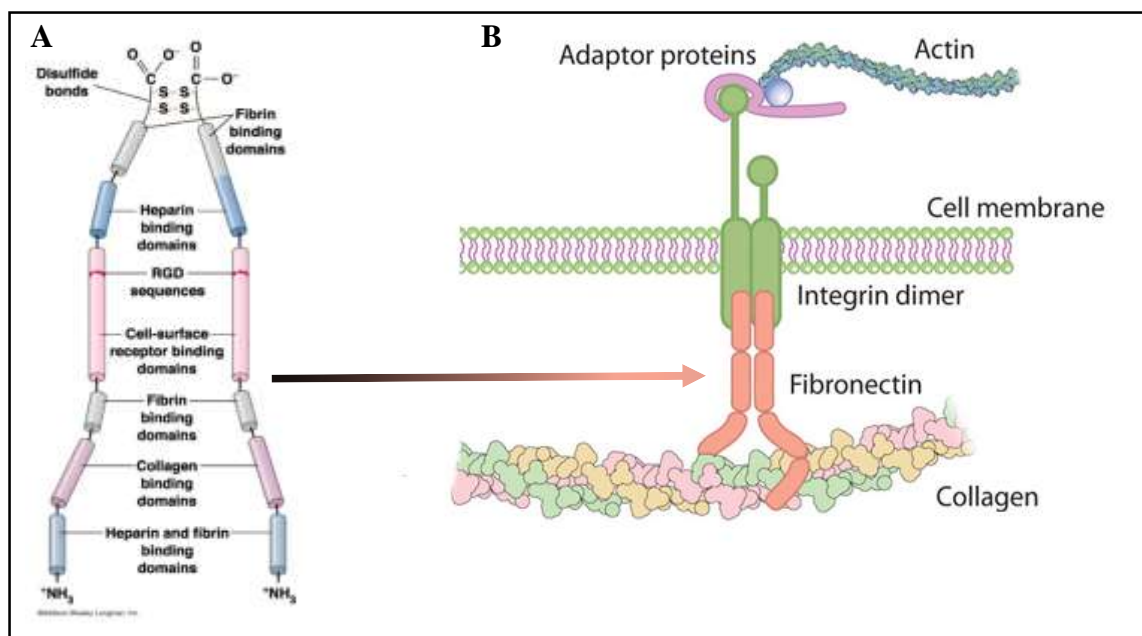


Figure 3 : Domaines d'interaction et organisation des domaines de la fibronectine (A). Représentation d'une interaction cellule-matrice via la fibronectine (B). D'après Addison Wesley Longman Inc, et Nature Education 2010

ces dernières se fait *via* un module tripeptidique « RGD » (arginyl-glycyl-aspartate) reconnu par les intégrines. La fibronectine est également capable de se lier aux fibrines (rôle dans l'hémostase) et aux glycosaminoglycanes (Mao & Schwarzbauer 2005). Parmi les glycoprotéines les plus étudiées se trouvent également les laminines, les tenascines, le nidogène et les thrombospondines possédant leurs propres modules d'interaction. Les glycoprotéines matricielles ont des rôles physiologiques importants, notamment dans la croissance cellulaire en emprisonnant certains facteurs de croissance qui seront relargués lors de la protéolyse de celles-ci (Pankov & Yamada 2002). Des rôles ont également été montrés lors de l'embryogenèse (George et al. 1993).

4. Les molécules élastiques

L'élasticité est principalement assurée par les élastines, et dans une moindre mesure les fibrillines qui sont des protéines aux propriétés élastiques sécrétées jusqu'à la fin de la croissance de l'organisme par les fibroblastes. L'origine de l'élasticité de ces protéines réside dans leur séquence primaire. En effet, elles sont composées d'acides aminés hydrophobes à plus de 80% qui vont former au repos des clusters hydrophobes, repliant ainsi la protéine. Les élastines sont également reliées entre elles par des liaisons desmosines ou isodesmosines. Lors de l'étirement des molécules, les interactions hydrophobes sont rompues, les protéines sont alors dépliées augmentant ainsi la longueur. Lorsque plus aucune force de tension n'est appliquée, les protéines vont revenir à un état thermodynamiquement favorable et reformer des interactions au niveau de leurs acides aminés hydrophobes ce qui va raccourcir la macromolécule (Figure 4).

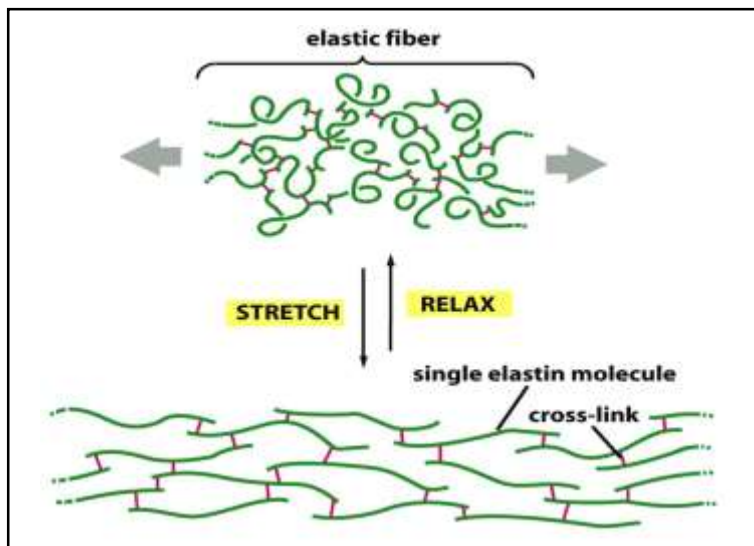


Figure 4 : Représentation d'un réseau d'élastine au repos (haut) et étiré (bas). D'après Molecular Biology of the Cell 5th edition.

Ces protéines sont retrouvées dans les matrices sujettes aux forces de tension comme au niveau de la peau, des poumons, des vaisseaux sanguins, des tissus conjonctifs, des tendons... Les molécules d'élastine sont également intimement liées aux fibres de collagènes dans la matrice extracellulaire et peuvent représenter plus de la moitié en masse du tissu sec (e.g. l'aorte) (Wise & Weiss 2009).

5. Les protéoglycans (PG)

La dernière grande famille composant les matrices extracellulaires est celle des protéoglycans (PG). Ce sont des protéines particulières sur lesquelles sont liées de manière covalente une ou plusieurs chaînes de sucres non-ramifiées appelées glycosaminoglycans

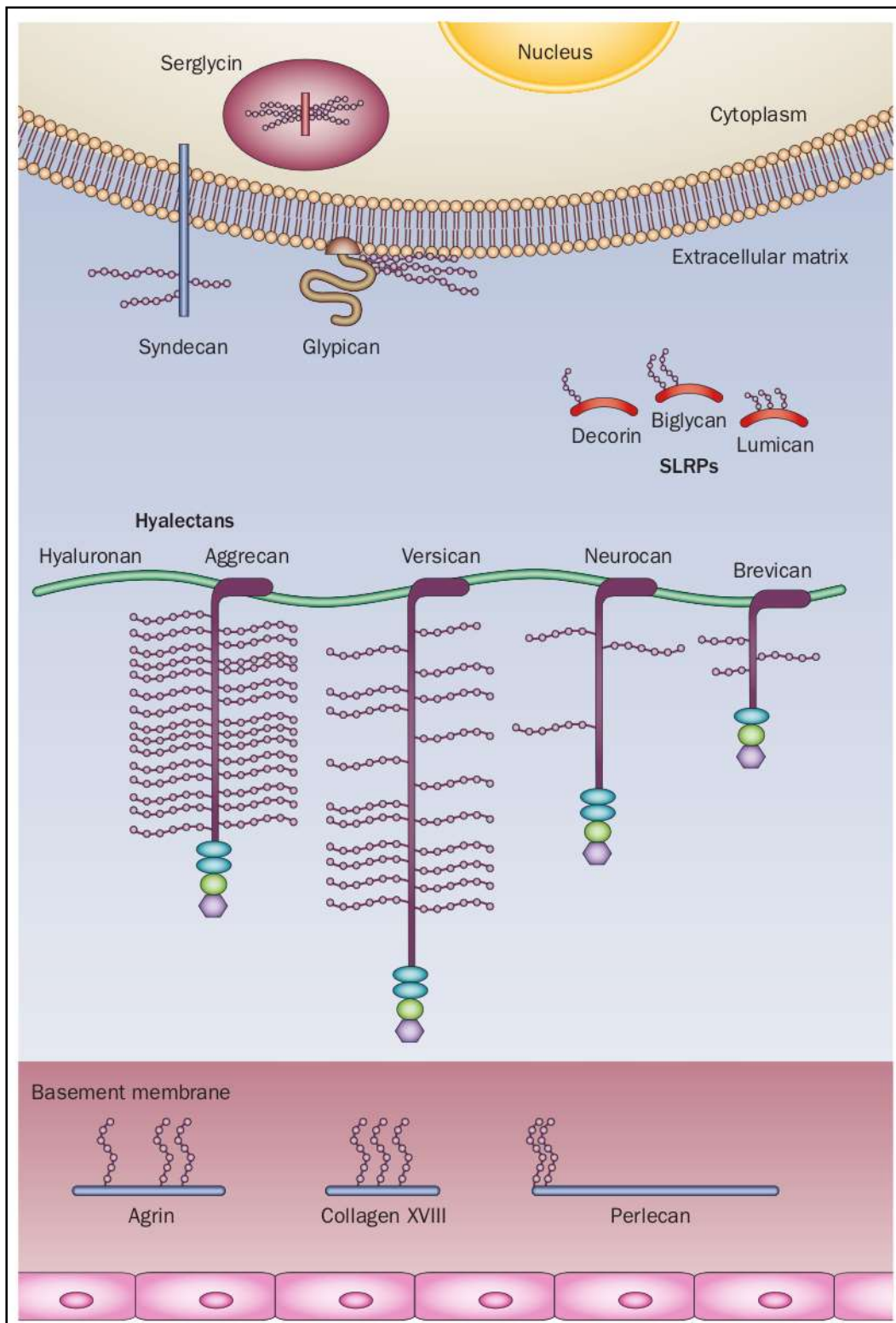


Figure 5 : Principaux protéoglycans cellulaires et matriciels et classification selon leur protéine cœur, leur localisation et leur structure. SLRP : Small Leucine Rich Proteoglycan. Les chaînes de GAG sont représentées en collier de perles. D'après (Edwards 2012).

(GAG) (Figure 5). Ces chaînes possèdent un certain motif disaccharidique répété selon leur partie protéique (appelée « protéine cœur ») mais sont généralement toutes composées d'une hexosamine liée à un acide uronique (sauf pour les kératanes sulfates où l'uronate est remplacé par un galactose). Parmi les GAG, on retrouve les Chondroïtines Sulfates (CS), Dermatanes Sulfates (DS), Héparanes Sulfates (HS), Héparine (Hp), Kératanes Sulfates (KS) et l'Acide Hyaluronique (AH). Ce dernier GAG a la particularité de n'être jamais lié covalamment à une protéine cœur.

De par leur nature hydrophile, les PG sécrétés avec leurs longues chaînes de GAG vont capter une grande quantité d'eau et d'ions au niveau de la matrice, assurant ainsi le remplissage des espaces et leur lubrification. Les GAG et protéines cœur des PG peuvent former des associations intermoléculaires entre eux ainsi qu'avec d'autres protéines de la MEC. Ces interactions ainsi que le pouvoir rhéologique des GAG vont être responsables de la résistance à la compression. Cela joue un rôle essentiel au niveau des cartilages comme dans les résistances mécaniques d'autres tissus comme la peau, les tendons...

Au cours de leur polymérisation, les chaînes de GAG vont pouvoir être modifiées principalement par des sulfatations et épimérisation à certaines positions du motif disaccharidique lui conférant une certaine structure. Cela va permettre d'interagir avec des molécules de type cytokines, facteurs de croissances (Tufvesson & Westergren-Thorsson 2002)... et de créer un réservoir de protéines au niveau des tissus sur lesquels elles vont pouvoir agir. Il existe plus d'une trentaine de protéoglycanes recensés dans la matrice classés selon leur localisation, leur protéine cœur et leur(s) chaîne(s) de GAG, ainsi que par leur séquence de reconnaissance d'autres molécules (glycoprotéines, GAG, protéines...) (Iozzo & Murdoch 1996). Du fait de leur très grande diversité et leur capacité d'interaction, les PG jouent des rôles variés dans le développement, lors de réponses immunitaires (Iozzo 1998)... et également dans des processus pathologiques tel que dans le cancer, l'adsorption de pathogènes (Iozzo & Sanderson 2011) (Guo et al. 1998)...

Les PG sont également retrouvés à la surface de toutes les cellules de mammifères ce qui en fait un acteur essentiel dans les nombreux processus physiologiques et pathologiques comme vu précédemment. L'étude de cette famille de molécules est devenue un sujet majeur pour la compréhension des phénomènes dans lesquels ils sont impliqués.

B. Les Protéoglycanes

1. Historique

a) Découverte du premier GAG : histoire autour du « chondroitinschwefelsäure »

En 1884, Krukenberg isole une molécule à partir de cartilage hyalin (principalement présent au niveau des articulations) par extraction en milieu alcalin dilué à température ambiante : c'est la première préparation d'un GAG (Krukenberg 1884).

En 1891, Schmiedeberg nomme cette préparation le *Chondroitinschwefelsäure* (acide sulfurique de chondroïne) de *chondros* signifiant le cartilage, ce qui va devenir « chondroïtine sulfate » par la suite. Il en déduit également la composition, à savoir : de l'acétate, sulfate, acide glucuronique et hexosamine (Schmiedeberg 1891).

Les recherches se poursuivent et ce n'est qu'en 1914 que la structure de l'hexosamine est déterminée par Levene et Laforge, d'abord appelée chondrosamine puis renommée galactosamine par la suite (Levene & LaForge 1914).

La molécule de chondroïtine sulfate est connue pour être un enchainement disaccharidique (appelé chondrosine), et en 1941 Levene et Laforge propose une structure erronée de ce motif, en proposant l'acide hexuronique comme porteur de l'extrémité réductrice (Levene & LaForge 1941). Ce n'est que 10 ans plus tard que Masamune *et al* corrigent cette structure en proposant la 1^{ère} structure correcte avec la galactosamine porteuse de l'extrémité réductrice (Masamune et al. 1951).

Davidson et Meyer déterminent en 1955 la liaison dans le motif disaccharidique [GlcA β 1-3GalNAc] (Davidson & Meyer 1955) puis entre les unités disaccharidiques [-4GlcA β 1-3GalNAc1 β -] en 1956 (Meyer et al. 1956).

Par la suite, ce sont les différentes positions des sulfatations qui ont été déterminées par des analyses infrarouges (Mathews 1958) (Hoffman et al. 1958).

b) Découverte des protéines cœur du PG

Du fait des conditions d'extraction des chondroïtines sulfates (en milieu alcalin), la structure proposée du cartilage était alors faite d'interactions ioniques entre les CS et les collagènes (Meyer & Smyth 1937). C'est en 1954 que Shatton et Schubert mettent en évidence la présence de protéines non-collagènes ne pouvant être séparées des CS en solution aqueuse à pH neutre. Cette protéine est alors appelée « protéine polysaccharide » et est en fait la première mise en évidence d'un PG qui s'avère être l'aggrécane (Shatton & Schubert 1954).

En 1958, Helen Muir apporte une nouvelle preuve de l'existence des PG en démontrant la présence d'acides aminés toujours liés à la chaîne de GAG après une digestion protéique à la papaïne (Muir 1958). Elle montre également que la sérine est le seul acide aminé ayant un ratio molaire de plus de 1 avec la chaîne de GAG, indiquant donc que les chaînes de GAG sont liées à ce résidu sérine.

2. Les protéines cœur

a) Classification

Les protéoglycanes peuvent être classés selon 4 critères en fonction de leur localisation au niveau de la cellule et de l'organisme (PG de la MEC, membranaires, intracellulaires ou circulants) :

(1) Les PG de la matrice extracellulaire

Dans cette famille, on peut distinguer 3 sous-ensembles selon leurs motifs protéiques :

(a) Les SLRP (*Small Leucine Rich Proteoglycans*)

Cette sous-famille est constituée de PG ayant une protéine cœur d'environ 40kDa avec une organisation en domaine au niveau de leur séquence primaire. Les SLRP possèdent des motifs riches en acide aminé Leucine appelés LRR (pour Leucine Rich Repeats) de séquence LxxLxLxxNxI (où L est généralement une leucine, N une asparagine, I une isoleucine, et x un acide aminé quelconque) (Figure 6), le tout formant un repliement tridimensionnel en « fer à

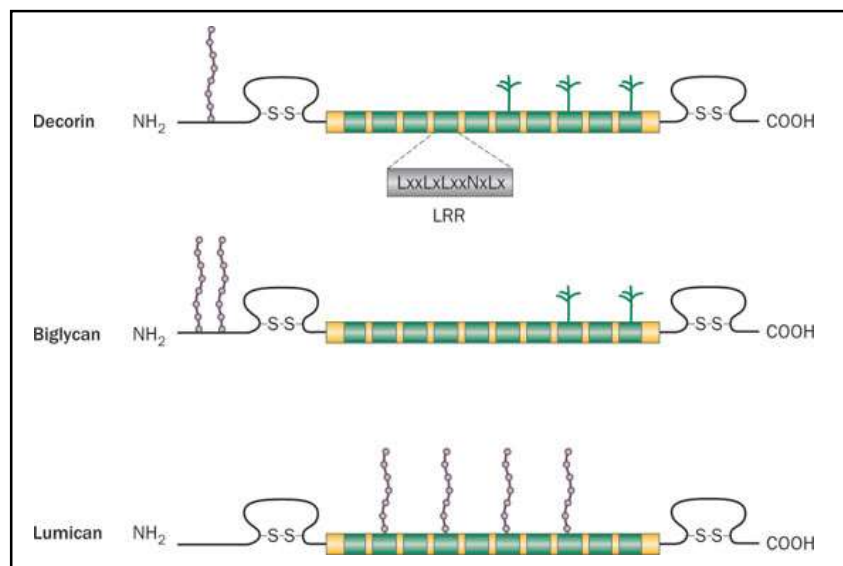


Figure 6 : Organisation structurale de trois différents SLRP. Les GAG (chondroïtine/dermatane/kératane sulfates) sont représentés en collier de perle. D'après (Edwards 2012).

cheval » (Kobe & Deisenhofer 1993) (Scott et al. 2006) (Figures 5 et 7). Au cours de l'évolution, 4 cystéines N-terminales et 2 en position C-terminale ont été conservées et forment des ponts disulfures, jouant un rôle structural important (Figure 6).

On retrouve principalement parmi les SLRP, le biglycan, la décorine, le lumican et la fibromoduline qui sont les plus étudiés (Figure 5). Des séquences consensus serylglycines sont retrouvées dans leur séquence primaire permettant la polymérisation de chaînes de chondroïtines ou dermatanes sulfates (CS/DS) au niveau de ces sérines. Des chaînes de kératane sulfate (KS) peuvent également y être liées ou au niveau de l'asparagine des LRR (Figure 6).

Les SLRP peuvent interagir au niveau de la matrice avec leur protéine cœur et *via* leur chaîne de GAG et former des interactions protéine-protéine ou protéine-sucre. Ainsi, ils peuvent interagir avec les collagènes de la MEC (Svensson et al. 1995) (Kalamajski & Oldberg 2009), comme avec des cytokines telles que le TGF- β (Kolb et al. 2001), TNF- α (Tufvesson & Westergren-Thorsson 2002), BMP-4 (Chen et al. 2004), PDGF (Nili et al. 2003)... participant à la régulation de leur fonction au niveau spatial et temporel. Les SLRP sont également capables

d'interagir directement avec les cellules *via* des récepteurs membranaires comme les *Toll-like receptors* (TLR) (Schaefer et al. 2005), l'EGFR (Iozzo et al. 1999) et donc de réguler certaines activités physiologiques (Figure 7).

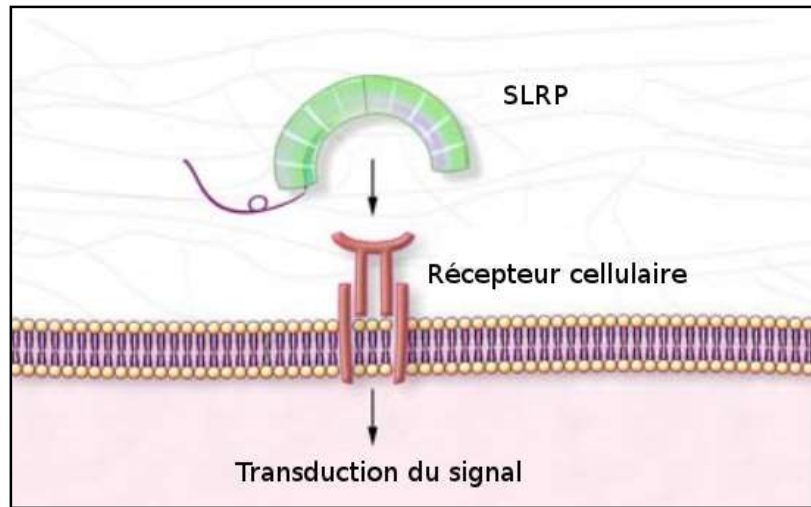


Figure 7 : Schéma d'interaction entre un SLRP et un récepteur cellulaire via la protéine cœur. D'après (Merline et al. 2009).

(b) Les hyalectans (ou lecticans)

Hyalectan est la contraction de « hyaluronane » et « lectine » car ces PG ont la propriété *via* leurs domaines *link* et lectine de type C d'interagir à la fois avec l'acide hyaluronique et les sucres de glycoprotéines en faisant intervenir un ion calcium (comme les ténascines). Les différents hyalectans ont une organisation protéique semblable. En effet, les domaines d'interaction *link* N-terminal et lectine C-terminal sont des domaines globulaires conservés entre les différents membres (Sandy et al. 1990) (Kohda et al. 1996) (Brissett & Perkins 1998). Les hyalectans se différencient entre eux par leur région centrale dont la séquence primaire et la longueur varient. C'est aussi au niveau de ce domaine que sont attachées les chaînes de GAG pour former le PG.

La famille des hyalectans est composée de 4 membres qui sont le versican, l'aggrécan, le neurocan et le brévican (Figure 5). Ce sont des PG à chondroïtine sulfate (CSPG) strict à l'exception de l'aggrécan sur lequel des chaînes de kératane sulfate sont aussi polymérisées. Leur localisation est aussi différente, bien que l'aggrécan et le versican soient assez ubiquitairement exprimés dans l'organisme, les neurocans et brévicans quant à eux sont principalement retrouvés au niveau du cerveau (Rauch et al. 2001) (Yamaguchi 1996). Ils jouent un rôle important dans les migrations cellulaires, dans la formation de motifs pendant le développement du système nerveux périphérique *etc.* L'aggrécan joue un rôle primordial également dans les cartilages de par sa très grande taille et sa capacité à remplir les espaces en occupant un large volume hydrodynamique.

(c) Les autres PG de la MEC

Ils ne constituent pas une famille à part entière car ils ne possèdent pas de motifs communs entre eux. On y retrouve 5 types de collagènes (dont le collagène XVIII précédemment détaillé) qui forment des molécules hybrides. Il y a également une isoforme de l'agrine qui joue des rôles importants dans le cerveau et l'immunité (Bezakova & Ruegg 2003), le bamacan au niveau des membranes basales et le perlécan également retrouvé dans les lames basales (Figure 5) et qui possède un grand nombre de partenaires protéiques comme des facteurs de croissance (Mongiat et al. 2000), intégrines, protéines de la matrice...et joue par voie de conséquence des rôles très variés, notamment pour la formation de ces lames basales (Behrens et al. 2012).

En effet, le perlécan est un PG à HS (HSPG) avec une masse moléculaire de plus de 800kDa ce qui en fait une des plus grosses protéines chez l'Homme. Il est constitué de 5 domaines indépendants notés de I à V (Figure 8). Le domaine I N-terminal possède 3 séquences d'ancrage aux chaînes d'HS, et ces dernières vont notamment interagir avec des facteurs de croissance et constituer un réservoir de protéines. Ce domaine est le seul qui soit

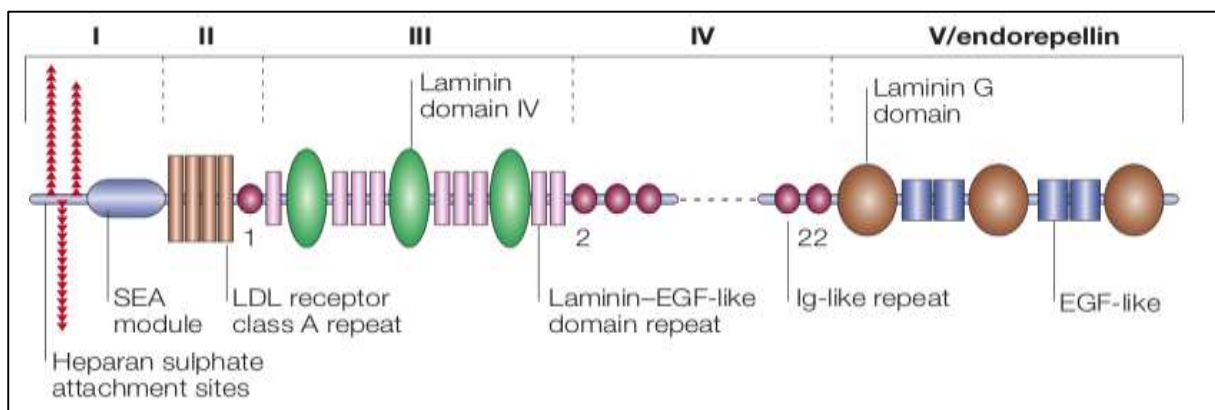


Figure 8 : Schéma d'organisation du perlécan. Les cinq domaines sont numérotés en chiffres romains de l'extrémité N-terminale à C-terminale. SEA module : Sperm protein, enterokinase, agrin module ; LDL : Low density lipoprotein ; EGF-like : Epidermal growth factor-like ; Ig-like : Immunoglobuline-like. D'après (Iozzo 2005).

unique dans le perlécan, les autres domaines partageant une très grande homologie de séquence avec d'autres protéines de la MEC (Murdoch et al. 1992). Le domaine II est capable d'interagir avec des lipoprotéines (comme le LDL) et les domaines III, IV et V via leurs différents motifs peuvent se lier aux intégrines cellulaires, à la MEC et à des facteurs de croissance. Les souris KO pour le gène codant le perlécan subissent une mort embryonnaire ou néonatale et présentent de nombreuses anomalies au niveau des cartilages, des os, dans le développement de certains organes (Iozzo 2005) (Knox & Whitelock 2006) (Farach-Carson et al. 2014)...

(2) Les PG membranaires

Les PG attachés à la membrane plasmique sont répartis en 4 familles (CD44, bétaglycan, syndécans et glypicans) :

(a) Le CD44

Le CD44 est codé par un seul gène mais le transcrit peut subir différents épissages aboutissant à plusieurs isoformes du CD44. Ce sont des protéines transmembranaires avec un domaine C-terminal cytosolique, un domaine transmembranaire et un domaine N-terminal extracellulaire. Ce dernier domaine varie en longueur selon son épissage. C'est au niveau de cet ectodomaine que sont retrouvées les chaînes de GAG adjointes : principalement des CS sauf pour une isoforme possédant préférentiellement des HS.

Le CD44 standard (CD44s) interagit avec l'acide hyaluronique de la MEC (Aruffo et al. 1990) (Lesley et al. 2000), mais toutes les isoformes n'en sont pas capables. Les CD44 interviennent dans de nombreux processus cellulaires (adhésion, survie, différenciation...) grâce aux multiples interactions possibles de molécules signal au niveau soit de la protéine cœur, soit de la chaîne de GAG (Ponta et al. 2003) (Figure 9). Ils sont exprimés dans de nombreux types cellulaires.

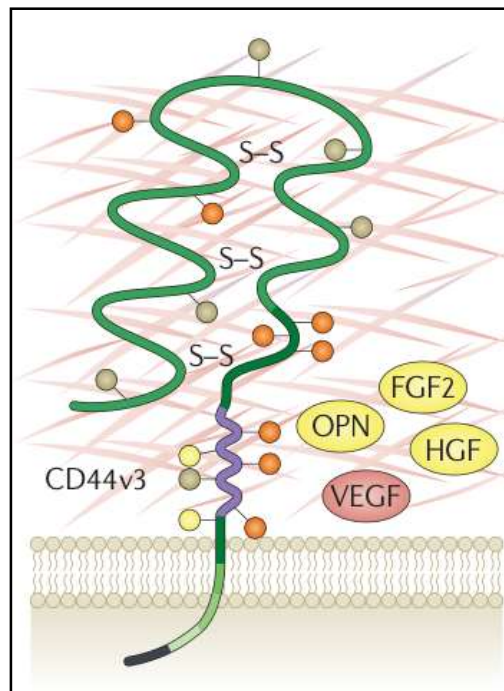


Figure 9 : Organisation d'une isoforme du CD44 (CD44v3) porteuse de deux chaînes de GAG (cercles jaunes). Les O-glycosylations sont représentées par des cercles orange, les N-glycosylations par des cercles verts. OPN : ostéopontine ; FGF2 : Fibroblast growth factor 2 ; HGF : Hepatocyte growth factor ; VEGF : Vascular endothelial growth factor. D'après (Zöller 2011).

(b) Le bétaglycan

Le bétaglycan a été découvert en premier lieu comme corécepteur du TGF β (Massagué 1985), cytokine d'intérêt dans les phénomènes de prolifération et différenciation cellulaire, puis comme PG à HS/CS principalement (Segarini & Seyedin 1988). Il possède un court domaine cytoplasmique, une hélice transmembranaire, et un long domaine N-terminal. L'interaction TGF β -bétaglycan ne déclenche pas directement de voie de signalisation, mais il sert principalement de réservoir pour cette cytokine et pour la présentation de cette dernière à d'autres récepteurs médiant la transduction du signal (Wang et al. 1991). Il est exprimé dans presque la totalité des membranes plasmiques et il existe également une forme soluble du bétaglycan, issue de son clivage par des protéases au niveau de la surface cellulaire, (phénomène de *shedding*) (Velasco-Loyden et al. 2004) qui est retrouvée dans les matrices et dans le sérum.

Les rôles du bétaglycan membranaire sont liés à ceux du TGF β . Il joue également un rôle suppresseur de tumeur important (Turley et al. 2007), et des dérégulations de son expression (généralement sa sous-expression) favorisent des phénomènes pathologiques comme certains cancers (Hempel et al. 2007).

(c) Les syndécans

Les syndécans forment une famille contenant 4 membres nommés syndécan-1 (syndécan), syndécan-2 (fibroglycan), syndécan-3 (N-syndécan) ou syndécan-4 (ryudocan/amphiglycan) selon l'ordre de clonage de leur ADNc. Ils sont codés par des gènes situés sur 4 chromosomes différents et exprimés de manière tissu-dépendante (Couchman 2010) mais possèdent des caractéristiques communes. En effet, tous les syndécans sont des PG transmembranaires de type I avec une courte queue cytoplasmique C-terminale, un domaine unique transmembranaire bien conservé et un domaine N-terminal extracellulaire

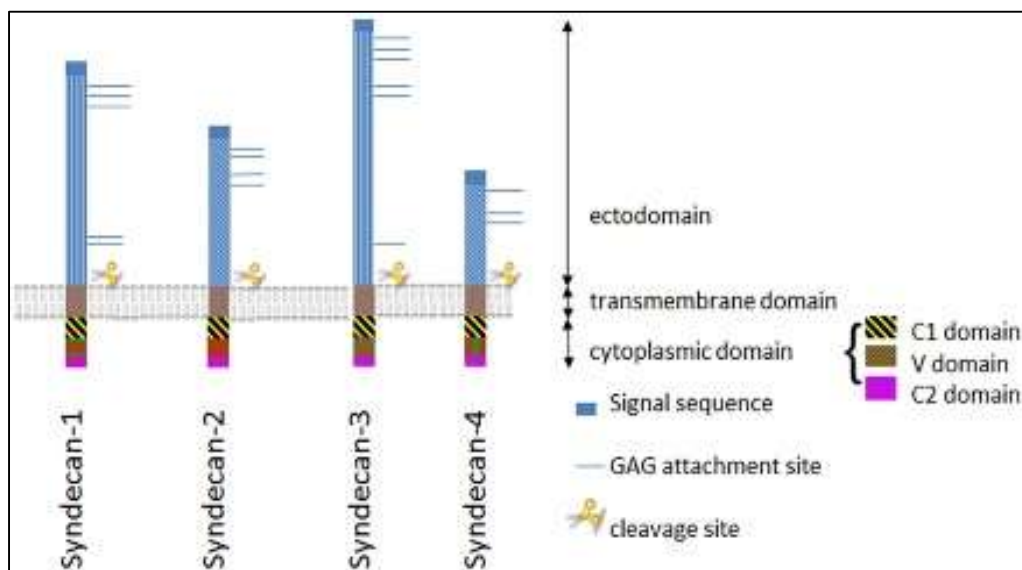
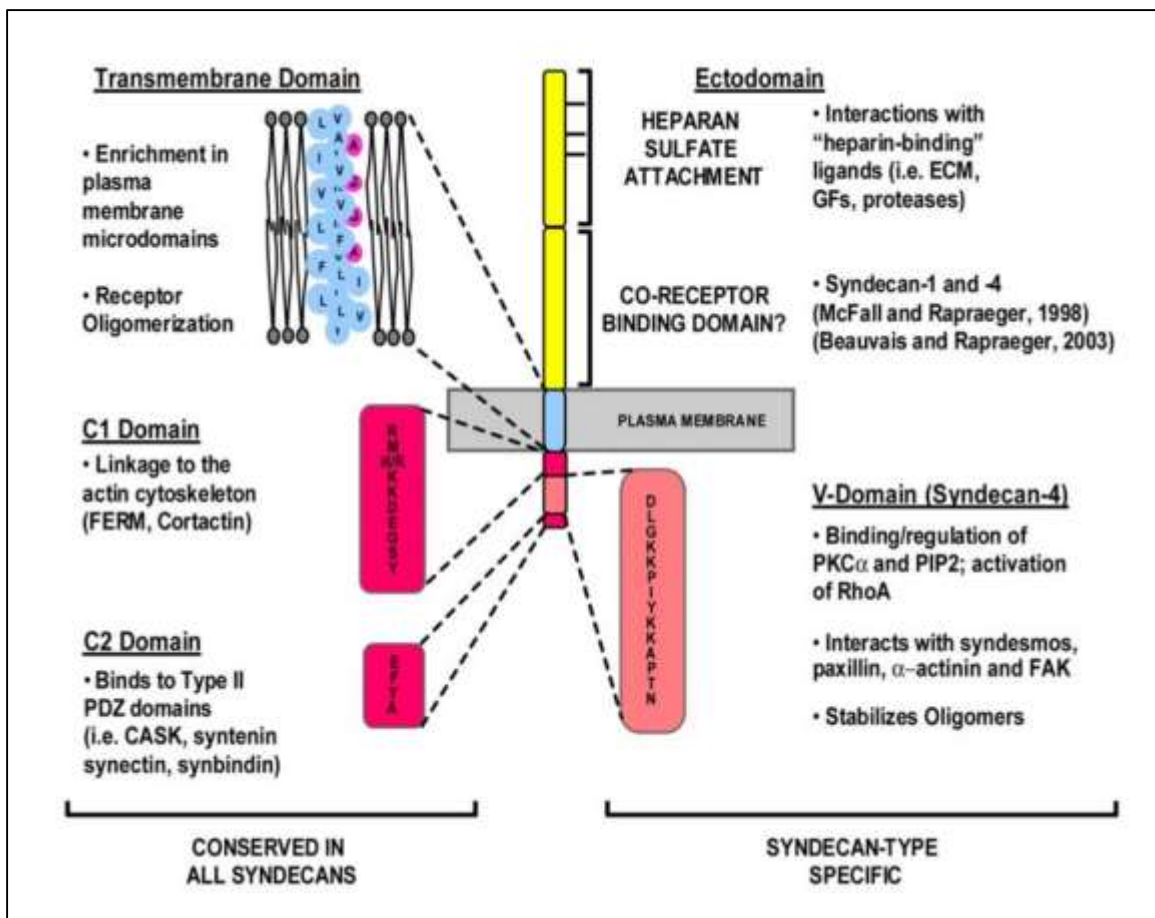


Figure 10 : Représentation schématique des différents membres de la famille des syndécans. D'après (Sztamári & Dobra 2013)

appelé ectodomaine qui est le moins conservé des 3 domaines entre les différents syndécans (Bernfield et al. 1992). Celui-ci varie en longueur et dans sa séquence primaire (10 à 20% d'homologie entre les 4 ectodomaines) mais des séquences peptidiques serylglycine sont conservées et forment les séquences consensus sur lesquelles les chaînes de GAG vont être greffées (Figure 10). Ce sont principalement des PG à HS (HSPG) même si les syndécans 1 et 3 sont porteurs également de quelques chaînes de CS (Couchman 2010). Les rôles des syndécans sont principalement liés à leurs chaînes d'HS que nous détaillerons plus tard. Brièvement, les HS du fait de leur très grande diversité structurale (plus importante que tous les autres GAG) et de leur nature polyanionique, peuvent interagir avec de nombreuses protéines et potentialiser leurs actions. En effet, ils concentrent les protéines à la surface cellulaire, forment des gradients de concentration (rôle essentiel dans l'immunité), favorise la dimérisation de protéines, les protègent de la protéolyse (Sadir et al. 2004), favorise les contacts cellule-matrice et cellule-cellule (Sarrazin et al. 2011)...



pouvant interagir avec des éléments de transduction de signaux intracellulaires et déclencher une réponse de la cellule (Beauvais & Rapraeger 2004) (Figure 11).

(d) Les glypicans

Chez l'Homme, les glypicans constituent une famille de 6 PG (Gpc-1 à -6) codés par 6 gènes différents. Leur homologie de séquence varie de 35 à 60% et ils gardent une organisation similaire. En effet, à leur extrémité C-terminale se situe une ancre glycosylphosphatidylinositol (GPI) qui est un glycolipide enchâssé dans le feuillet externe de la membrane plasmique, exposant donc la protéine au niveau extracellulaire. Cette extrémité C-terminale des glypicans d'environ 50 acides aminés est également hydrophobe et contient les séquences consensus serylglycines de liaison aux GAG. Ces chaînes de GAG sont des HS, faisant des glypicans des HSPG (à l'exception du glypican-5 capable d'avoir des chaînes de CS (Saunders et al. 1997)) dont les chaînes se situent proches de la membrane plasmique. Les glypicans possèdent leur séquence signal d'export à la surface membranaire du côté N-terminal et l'expression de ces protéoglycanes est spécifique du tissu. La protéine cœur est constituée de 14 cystéines conservées formant 7 ponts disulfures dans la protéine (Veugeliers et al. 1999). Ces ponts ont une importance dans la structure globulaire tridimensionnelle de la protéine (Figure 12).

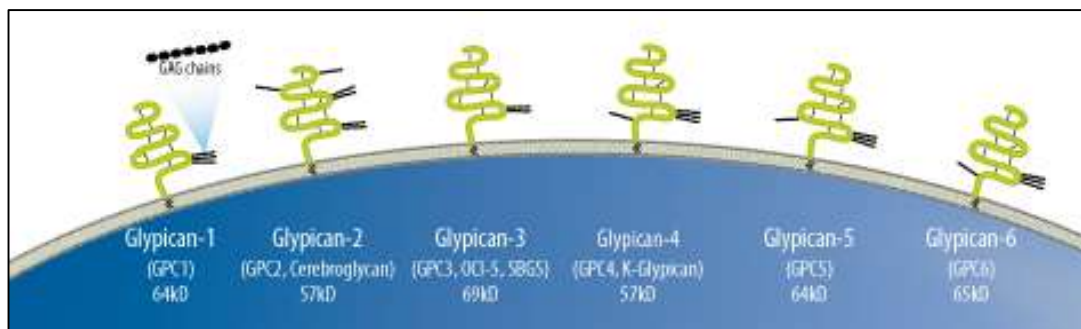


Figure 12 : Représentation des différents membres de la famille des glypicans.

L'ancre GPI des glypicans favorise leur localisation au niveau de rafts lipidiques (Mayor & Riezman 2004), et permet également un renouvellement rapide de la protéine. Dans les cellules polarisées, les glypicans (sauf Gpc-1) sont principalement retrouvés à la surface apicale de la cellule et dans une moindre mesure au niveau basolatéral. Les chaînes d'HS joueraient un rôle dans cette localisation cellulaire (Mertens et al. 1996).

Les glypicans peuvent interagir avec des morphogènes (Wnt, Shh...) et sont cruciaux au cours de l'embryogenèse, et lors de la croissance cellulaire de par leurs interactions avec les facteurs de croissance (Song et al. 1997) et également avec des récepteurs cellulaires. Les interactions entre HS, morphogènes et récepteurs sont favorisées par la localisation des chaînes d'HS situées proches de la membrane plasmique. Toutefois, les protéines cœurs jouent un rôle essentiel dans certaines régulations de morphogènes comme Wnt (Song et al. 2005), Shh (Capurro et al. 2008) (Kirkpatrick et al. 2006). Une expression anormale des glypicans est observée dans de nombreux cancers. C'est le cas du glypican-1 qui est

surexprimé dans le cancer du pancréas, du sein (Filmus et al. 2008)... alors que le glypican-3 est sous-exprimé lors de la progression tumorale (Lin et al. 1999), voire même réexprimé dans des tissus où il n'est pas exprimé physiologiquement (Filmus 2001).

(3) Les PG intracellulaires

Les PG sont en général extracellulaires. Ils ne sont retrouvés au sein des cellules que lors de leur synthèse et pendant leur recyclage. Toutefois, il existe un PG dont la localisation est intracellulaire : la serglycine. Elle est retrouvée stockée dans des granules de sécrétion des cellules hématopoïétiques et endothéliales (Kolset et al. 2004) et ce fut le premier PG dont la séquence protéique a été déterminée en 1988 puis cloné en 1989 (Ruoslahti 1989). Il fut nommé serglycine du fait de la présence de plusieurs motifs peptidiques (8) sérylglycines qui forment les points d'ancrages aux chaînes de GAG.

La serglycine est un PG avec une petite protéine cœur (15kDa) qui porte des chaînes de GAG très sulfatés. Selon le type cellulaire et sa localisation dans l'organisme la chaîne de GAG ne sera pas la même. Ainsi, dans les mastocytes des tissus conjonctifs la serglycine porte des chaînes de GAG de type héparine (Hp) qui possèdent le même squelette osidique que les HS mais étant beaucoup plus modifiées (voir structure Figure 27). Ce sont les seules cellules de l'organisme à posséder des chaînes d'héparine. Quant aux mastocytes des muqueuses, leur serglycine présente des chaînes de CS disulfatées de type –E (voir structure Figure 26). Les serglycines issues de monocytes et macrophages inactifs possèdent des chaînes de CS monosulfatées de type –A, mais lors de phénomènes inflammatoires et donc d'activation, des chaînes de CS-E sont retrouvées sécrétées par les macrophages. Les conditions d'inflammation seraient liées à l'augmentation de la sulfatation des chaînes de GAG portées par les serglycines (Uhlén-Hansen et al. 1989).

Le choix de chaînes d'Hp ou de CS dépend moins de la protéine cœur que de l'activité des différentes enzymes de synthèse des différents GAG présentes dans la cellule, de la concentration en sucres disponibles pour le métabolisme des chaînes *etc* (Prydz & Dalen 2000). Ces chaînes de GAG très modifiées sont capables d'interaction avec un très grand nombre de protéines telles que des protéases (élastase, carboxypeptidase, tryptase, granzymes...) (Humphries et al. 1999), l'histamine, des cytokines, des chimiokines pro-inflammatoires (Kolset et al. 1996) (Galli et al. 2005) (Figure 13). Les GAG de la serglycine jouent donc un rôle de rétention et de protection de ces médiateurs. Constitutivement ou lors d'activation et de stimulation du mastocyte, les granules de sécrétion sont libérés (Kulseth et al. 1999) et déclenchent le phénomène d'inflammation pour combattre l'infection par exemple en libérant les composés stockés et en interaction avec les GAG. De par sa force d'interaction avec de nombreuses molécules, la serglycine forme de gros complexes ce qui éviterait la diffusion des signaux pro-inflammatoires en dehors du site de l'inflammation (Kolset & Tveit 2008).

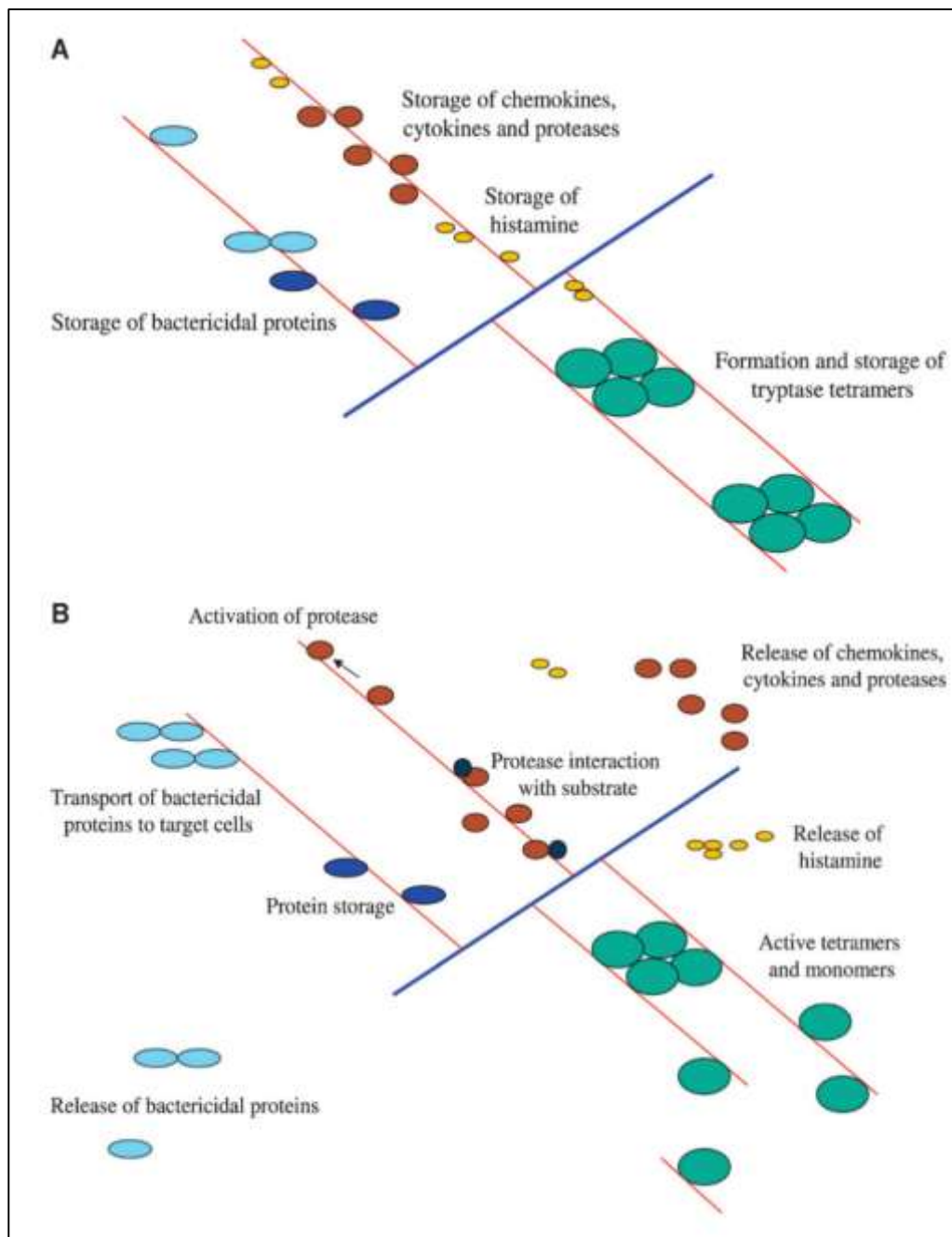


Figure 13 : Représentation schématique des rôles fonctionnels de la serglycine. A l'intérieur des cellules, la serglycine est importante pour la formation de granules et de stockage des protéines (A). Après sécrétion, la serglycine peut engager différentes activités majoritairement liées aux protéines relarguées dans l'environnement cellulaire et précédemment retenues par les chaînes de GAG (B). La protéine cœur est représentée en bleu et 4 des 8 chaînes de GAG sont représentées en rouge. D'après (Kolset & Tveit 2008)

La serglycine est donc un acteur fondamental dans les phénomènes d'immunité et de défense de par son rôle au cours des phénomènes inflammatoires. Les souris KO pour le gène codant la serglycine sont viables mais présentent des défauts au niveau des granulocytes et lors d'infection bactérienne, la virulence est augmentée par rapport à des souris sauvages (Abrink et al. 2004).

(4) Les PG circulants

Ils forment une famille de protéoglycanes retrouvés dans la circulation sanguine qui est composée de 4 membres principaux :

-La bikunine (également appelée urastine, ulinastatine, migrine, ASTI, UTI ou HI30) : elle possède une seule chaîne de CS ; sa protéine cœur est composée principalement de deux domaines inhibiteurs de protéase appelés domaines Kunitz, et posséderait des propriétés anti-inflammatoires (Matsuzaki et al. 2004).

-Le PG-100 : nommé ainsi du fait de son poids moléculaire ($\approx 100\text{kDa}$), le PG-100 possède une unique chaîne de CS (Schwarz et al. 1990). Il est sécrété par les macrophages et la forme protéolysée du PG (non-porteuse de la chaîne de CS) va jouer un rôle dans la survie, la prolifération et la différenciation de phagocytes (Partenheimer et al. 1995). La forme non-clivée du PG-100 (aussi appelé MCSF-1 pour *Macrophage Colony Stimulating Factor*) servirait de réservoir potentiel en facteurs de stimulation (CSF) inactifs potentiellement mobilisables.

-L'apolipoprotéine O (ApoO) : c'est le seul PG à ce jour possédant une fonction apolipoprotéique, c'est-à-dire de transport de molécules hydrophobes dans le sang. L'ApoO peut posséder une ou plusieurs chaînes de CS et pourrait donc jouer un rôle contre l'accumulation de lipides (Lamant et al. 2006).

-L'endocan (ou ESM-1 pour *endothelial cell-specific molecule 1*) : il est porteur d'une unique chaîne de DS (Bécharde et al. 2001). Il est exprimé par les cellules endothéliales et son fort niveau d'expression dans le sang est corrélé aux phénomènes de cancer.

Ces PG en tant que tels ont été découverts assez récemment, et même s'ils sont associés à différents événements physiologiques et pathologiques avérés, leurs rôles et leurs régulations sont aujourd'hui encore très mal connus.

Les principaux PG précédemment cités sont regroupés en annexe 1.

b) Synthèse et tri

La synthèse des protéoglycanes commence comme la synthèse de n'importe quelle protéine, à savoir que le gène codé par la protéine cœur est transcrit en un ARNm dans le noyau, et exporté dans le cytoplasme pour sa traduction par les ribosomes. La transcription des gènes codant pour les PG peut être influencée par des signaux extérieurs (Okuyama et al. 2013). Lors de la traduction du transcrit, les premiers acides aminés N-terminaux forment une séquence d'adressage au réticulum endoplasmique (RE) qui sont ensuite clivés au cours de l'élongation de la protéine au travers de la membrane du RE. Selon la séquence de la protéine cœur, celle-ci va pouvoir être une protéine transmembranaire de classe I (c'est-à-dire des protéines avec une unique hélice transmembranaire et dont la partie C-terminale est cytosolique, c'est le cas des syndécans, CD44...) ou une protéine libre destinée à la sécrétion (décorine, aggrecan...). Dans le cas des glypicans qui ont une ancre GPI, leur précurseur synthétisé est une protéine transmembranaire de type I avec une séquence de reconnaissance

dans son domaine N-terminal par la GPI transamidase catalysant la réaction de glypiation des glypicans.

Au niveau du RE et de l'appareil de Golgi les protéoglycanes naissants vont subir des modifications post-traductionnelles comme des *N*-glycosylations servant au bon repliement des protéines et à leur adressage vers des compartiments plus en aval dans la voie de biosynthèse (Scheiffele et al. 1995). Des *O*-glycosylations et/ou *N*-glycosylations (c'est à dire la polymérisation de la chaîne de GAG) vont également avoir lieu plus tardivement lors du trafic du protéoglycane jusqu'à sa destination finale (Vertel et al. 1993). La biosynthèse de cette dernière sera détaillée dans la prochaine partie.

Sans autre signal de rétention, les PG sont constitutivement destinés à la voie de sécrétion. Dans le cas de la serglycine, le tri serait fait au niveau du réseau *trans*-golgien vers des granules immatures. C'est la composition de ce granule qui empêcherait son excrétion ainsi que l'absence de stimulus extérieur. Toutefois, le tri de la serglycine dans ces vésicules particulières est encore mal connu bien que différents modèles soient proposés dans la littérature (Arvan & Castle 1998) (Wernersson & Pejler 2014).

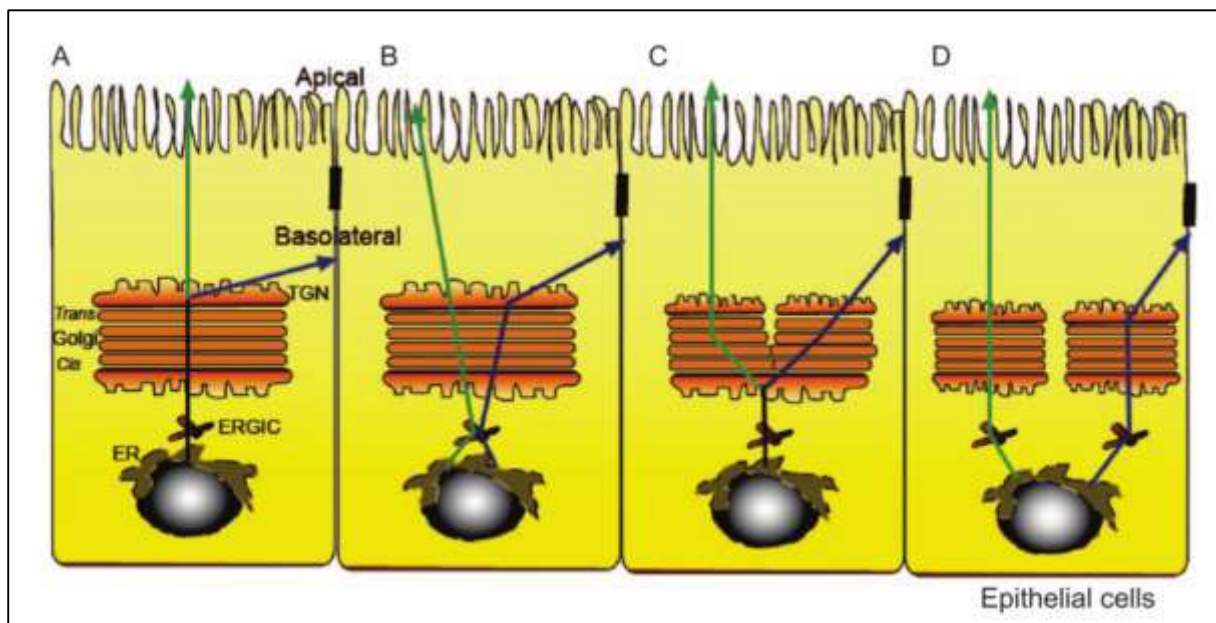


Figure 14 : Différentes hypothèses de voie d'adressage des protéines au travers du reticulum endoplasmic (ER) pour leur expression à une certaine polarité de la cellule épithéliale. Le tri entre pôle apical ou basolatéral est fait au niveau du réseau transgolgien (TGN) (A). La séparation des molécules selon leur lieu d'expression est réalisée au niveau du RE et tout au long de l'acheminement au travers des mêmes compartiments du ER (ERGIC) et de l'appareil de Golgi (B). La séparation des protéines est réalisée dans l'appareil cistolgien et adresse les molécules dans différents compartiments de l'appareil de Golgi (C). Les molécules transitent selon deux réseaux distincts (ER-ERGIC-appareil de Golgi) en fonction de la polarité cellulaire où elles doivent être exprimées (D). D'après (Prydz et al. 2008).

Enfin, un tri supplémentaire des PG est nécessaire dans les cellules polarisées n'exprimant pas les mêmes molécules selon leur côté apical ou basolatéral. La destination du PG semble être aussi bien liée à sa protéine cœur qu'aux chaînes de GAG qui lui sont attachées. En effet dans le cas du syndécan-1, la queue C-terminale cytoplasmique est essentielle pour son adressage du côté basolatéral. Lorsque celle-ci est tronquée suite à une délétion, il n'y a plus de signal spécifique pour son adressage et le PG est retrouvé sur toute

la surface cellulaire (Miettinen et al. 1994). La structure du GAG ainsi que ses modifications sont également porteuses d'information pour la localisation du PG. Il a été montré que les PG porteurs de chaînes HS et CS-C (C6S, chondroïtine 6-sulfate) étaient dirigés surtout du côté basolatéral (Vuong et al. 2006) (par exemple le perlécan, l'agrine, les glypicans...) alors que les PG à chaîne CS-A (C4S, chondroïtine 4-sulfate) étaient retrouvés du côté apical et basolatéral (Chihara et al. 1984). Le pH des vésicules jouerait également un rôle dans la polarisation de la sécrétion (Caplan et al. 1987). Bien que différentes preuves d'adressage aient pu être découvertes, les mécanismes de transport actif des vésicules vers la bonne localisation cellulaire sont encore flous (Matter 2000) (Figure 14).

c) Le catabolisme des PG

Afin d'assurer pleinement leurs rôles, les matrices extracellulaires ainsi que les surfaces cellulaires avec lesquelles elles peuvent interagir ont besoin de se renouveler. Les PG qui en sont des constituants majeurs ont besoin d'être recyclés en permanence, et cela tout au long de la vie d'un individu. La dégradation peut être enzyme-dépendante ou enzyme-indépendante. Dans cette dernière voie de catabolisme, ce sont des espèces réactives de l'oxygène (ROS) qui vont pouvoir altérer à la fois les protéines cœurs et les chaînes de GAG. Toutefois cette voie n'est pas une voie de dégradation totale des PG et est souvent liée aux phénomènes d'inflammation. La voie enzymatique, quant à elle, peut s'exprimer à différents niveaux. Un premier niveau est celui de l'expression de ces enzymes de dégradation au niveau extracellulaire. Elles vont être responsables du *shedding* des PG membranaires (Figure 15),

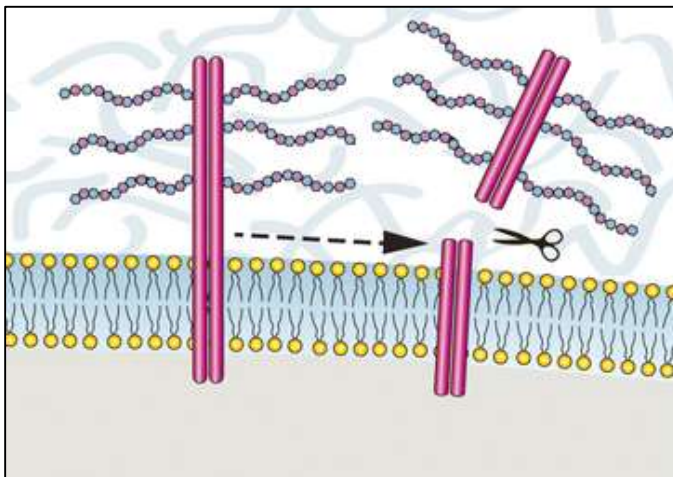


Figure 15 : Shedding d'un protéoglycane transmembranaire.
D'après (Strand et al. 2013).

c'est-à-dire de leur digestion en un fragment soluble relargué (pouvant jouer un rôle physiologique (Chen et al. 2009)), et en un fragment membranaire. Les protéases peuvent également fragmenter les PG de la MEC. Ces PG ainsi relargués vont pouvoir se retrouver dans le système circulatoire. Dans ce cas, ils vont être captés au niveau du foie par des cellules endothéliales qui vont les internaliser. Les autres PG peuvent être internalisés également par

pinocytose ou par endocytose *via* des récepteurs de GAG comme le CD44 par exemple. Une fois dans les endosomes, les PG vont être dirigés et dégradés dans les lysosomes exprimant les protéases (famille des cathepsines), les sulfatases et glycosidases requises pour une digestion totale. Quasiment toutes les cellules sont capables de cataboliser les PG car elles expriment les enzymes intracellulaires nécessaires. Les mutations de ces enzymes de dégradation sont associées à des maladies lysosomales de type mucopolysaccharidose où les PG et surtout les GAG sont accumulés (Bie et al. 2013). Dans le processus physiologique après

digestion, les protéines cœurs et chaînes de GAG sont réduites à leurs unités les plus simples qui seront réutilisés par la cellule *de novo*.

Les enzymes de dégradation des PG peuvent se classer dans 4 familles différentes selon leur site actif. On retrouve ainsi les sérine-protéases (les plus abondantes), les métalloprotéases (les plus importantes dans la MEC (Woessner 1994)), les aspartate-protéases et les cystéine-protéases. Ces deux dernières familles ont une activité à pH acide et sont donc actives dans les lysosomes tandis que les 2 premières catalysent leur réaction à pH neutre et ont donc une action au niveau de la MEC. Les protéines cœurs possèdent des sites de reconnaissance à ces différentes protéases qui sont assez spécifiques. L'action de ces protéases est contrôlée par l'expression d'une forme inactive de celles-ci. L'activation de cette dernière prend place au niveau des sites nécessitant leurs actions. Un bon équilibre de recyclage entre renouvellement des PG et leur dégradation est important afin de ne pas altérer les MEC dans lesquelles ils sont impliqués et qu'ils puissent exercer pleinement leurs fonctions, notamment de corécepteur (Elfenbein & Simons 2013). Les PG sont renouvelés en permanence dans l'organisme à des vitesses différentes liées à leur localisation (Yanagishita & Hascall 1992) (Figure 16).

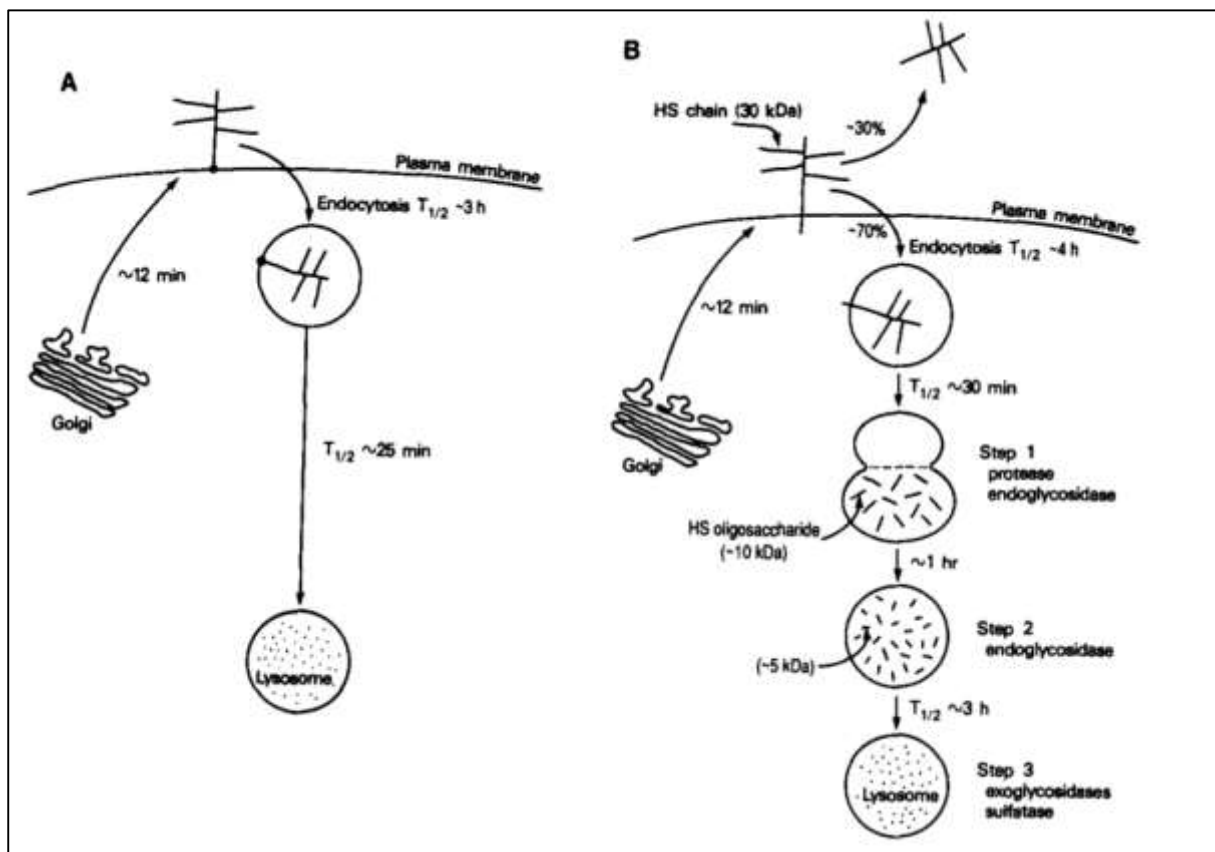


Figure 16 : Schémas modèle de métabolisme de protéoglycanes à ancre glycosylphosphatidylinositol (GPI) (A) et transmembranaires (B). D'après (Yanagishita & Hascall 1992).

C. Les GAG et leur métabolisme

1. Généralités

Les GAG forment une famille de 4 membres (acide hyaluronique, kératanes sulfates, chondroïtines/dermatanes sulfates, héparanes sulfates/héparine) possédant des points communs et des différences, même à l'intérieur d'une même famille.

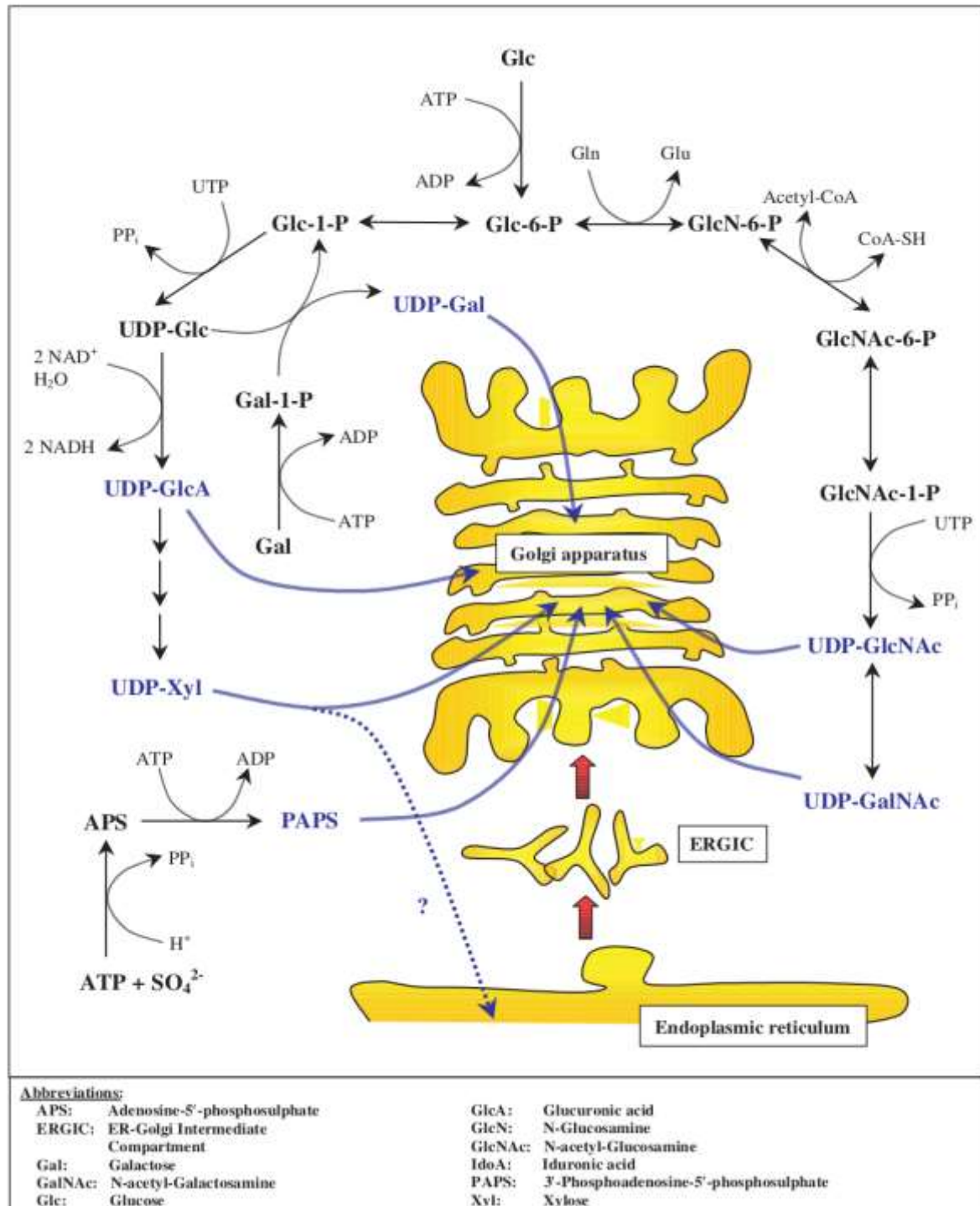


Figure 17 : Voies de synthèse des différents UDP-sucre et du 3'-phosphoadénosine 5'-phosphosulfate (PAPS) pour la formation et sulfatation des chaînes de GAG. Les différents éléments sont activement importés dans la lumière des différents compartiments par des transporteurs spécifiques. D'après (Prydz & Dalen 2000).

En effet, les GAG sont des chaînes polysaccharidiques non-ramifiées composées de la répétition d'une unité disaccharidique. La nature des oses composant cette unité ainsi que la nature de leurs liaisons définissent la famille du GAG. Mis à part l'acide hyaluronique, la synthèse des GAG a lieu principalement au niveau de l'appareil de Golgi. C'est également dans cet organe que les chaînes de GAG vont subir différentes modifications au cours de leur polymérisation. Tous les oses entrant dans la synthèse des chaînes sont activés par l'ajout d'une molécule d'UDP formant un UDP-sucre. Cette synthèse a lieu dans le cytoplasme, et ils peuvent être activement importés dans le reticulum endoplasmique (RE) et dans l'appareil de Golgi par des contre-transporteurs (Figure 17). Les groupements sulfates qui vont modifier les chaînes de GAG proviennent d'un donneur universel, le PAPS (3'-phosphoadénosine 5'-phosphosulfate), également importé activement dans les organites (Abeijon et al. 1997). L'efficacité de la polymérisation ainsi que les modifications des chaînes de GAG dépendent de la cellule et de son état physiologique ainsi que du tissu dans lequel elles sont produites.

La structure des GAG est essentielle dans les nombreux rôles qu'ils jouent au niveau des membranes plasmiques et de la matrice extracellulaire et est finement contrôlée. Les GAG sont renouvelés en permanence afin de répondre au mieux aux besoins des tissus de l'organisme dans lesquels ils se trouvent.

2. L'acide hyaluronique (HA)

L'HA a été découvert en 1934 par Meyer et Palmer (Meyer & Palmer 1934) et isolé à partir de l'humeur vitrée de l'œil d'où le nom d'acide hyaluronique (*hyalos* signifiant la vitrée), aussi appelée hyaluronane ou hyaluronate. C'est une molécule largement distribuée dans le monde du vivant car elle est retrouvée également dans des capsules bactériennes (*e.g* chez *Streptococcus pyogenes*) (Kendall et al. 1937). Des gènes codant les protéines de synthèse des HA (les HA synthétases ou HAS) ont été retrouvés aussi dans le monde viral (DeAngelis et al. 1997).

Comme vu précédemment, les HA ne sont jamais reliés à une protéine cœur et ne sont donc pas des PG *stricto sensu*. Les chaînes d'HA ont une masse moléculaire très importante et un fort pouvoir rhéologique leur conférant des propriétés hydrodynamiques remarquables, essentielles au niveau des articulations. Les HA jouent un rôle dans l'homéostasie et dans l'intégrité biomécanique, dans les mouvements cellulaires notamment au cours de l'embryogenèse, lors de la cicatrisation, mais aussi dans des phénomènes d'inflammation et de cancérogenèse.

a) Synthèse

L'acide hyaluronique (HA) est le seul glycosaminoglycane non lié à une protéine cœur, mais il diffère sous de nombreux autres aspects. En effet, le lieu de la synthèse de la chaîne ne se fait pas au niveau de l'appareil de Golgi comme nous le verrons pour les autres GAG

mais au niveau de la membrane plasmique *via* les acides hyaluroniques synthétases (HAS) I, II et III (Philipson et al. 1985).

Les HAS sont des enzymes possédant 7 domaines transmembranaires qui catalysent la formation des liaisons entre les acétyl-glucosamines (GlcNAc) et les acides glucuroniques (GlcA) ainsi qu'entre les acides glucuroniques et les acétyl-glucosamines qui composent le GAG (DeAngelis & Weigel 1994). Au cours de l'élongation de la chaîne, l'HA est relargué dans le milieu extracellulaire (Prehm 1984) par des transporteurs (Schulz et al. 2007). Ce mode de synthèse unique parmi les autres macromolécules permet de synthétiser de très longues chaînes d'HA sans les concentrer dans un compartiment où la viscosité induite pourrait être délétère. Les 3 isoformes d'HAS possèdent des propriétés différentes, notamment par les longueurs de chaînes d'HA qu'elles synthétisent de 200 à 2000kDa pour les HAS I et III (jusqu'à 5000 unités disaccharidiques) et plus de 2000kDa pour les HA synthétisés par l'HAS II (Itano et al. 1999).

Lors de la synthèse, les HAS interagissent avec la forme nucléotidique active des sucres qui composent l'HA (l'uridine diphosphate *N*-acétylglucosamine ou UDP-GlcNAc et l'uridine diphosphate glucuronate ou UDP-GlcA) par un mécanisme à 2 sites de fixation. Quand les 2 sucres sont fixés, l'enzyme catalyse la formation d'un disaccharide en faisant interagir l'extrémité réductrice du sucre en position non-réductrice avec le sucre en position réductrice. Cette réaction libère l'UDP du sucre en position non-réductrice. Le sucre en position réductrice lui est toujours lié à l'enzyme et à son UDP. La réaction libère un site de fixation où un nouveau UDP-sucre va pouvoir venir se fixer. La néo-chaîne d'HA va être liée à l'UDP-monosaccharide *via* son côté réducteur relargant ainsi son UDP. Le site vacant sur l'HAS va ainsi pouvoir interagir avec un nouveau sucre et la synthèse de la chaîne va pouvoir se poursuivre par l'ajout successif de GlcA et GlcNAc. Chaque site est spécifique d'un monosaccharide (Figure 18) (Prehm 1983). L'étape d'initiation de la polymérisation de la chaîne d'HA débiterait par la

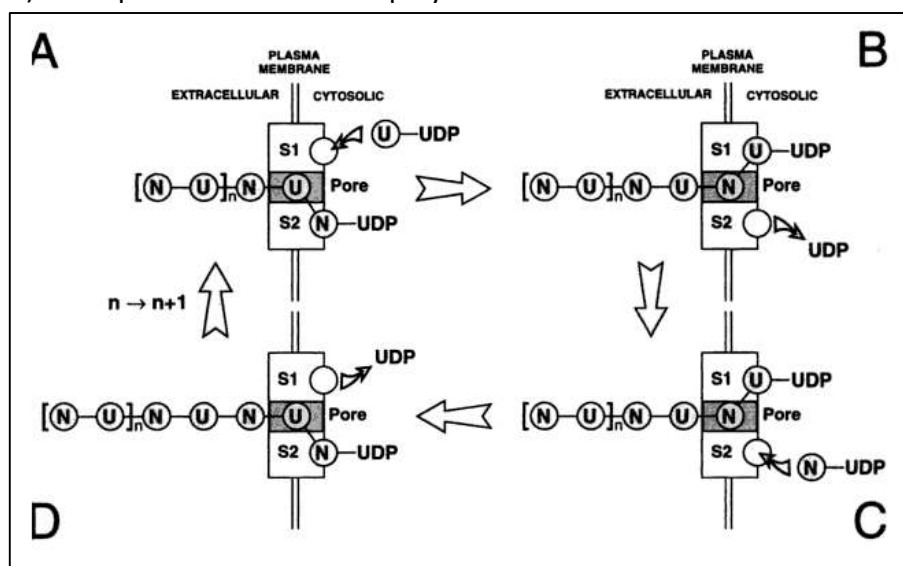


Figure 18 : Modèle à deux sites pour la synthèse de l'acide hyaluronique. Une chaîne de HA avec n disaccharides est montrée schématiquement (A). Un cycle (ABCD) ajoute un disaccharide à la chaîne de départ en cours de polymérisation à partir de deux UDP-sucres substrats.

fixation des précurseurs au niveau de l'HAS et ne se différencierait pas donc de l'étape d'élongation. L'étape de terminaison de la chaîne d'HA est également mal connue. Toutefois, l'expression des 3 isoformes de l'HAS est finement régulée ce qui pourrait s'expliquer par le fait que chaque enzyme est codée sur un chromosome différent (Spicer et al. 1997) (Pienimäki et al. 2001) (Karvinen et al. 2003).

b) Structure

La structure du motif disaccharidique [-4GlcA β 1-3GlcNAc β 1-] a été déterminée dans les années 1950 (Weissmann & Meyer 1954) (Figure 19). Du fait de sa synthèse au niveau de la membrane plasmique et de son extrusion au cours de celle-ci, le polysaccharide n'est pas modifié et la chaîne

native est donc sécrétée telle quelle dans le milieu extracellulaire. Une fois sécrétées, les chaînes d'HA vont se replier de manière aléatoire en formant des « pelotes » (on parle de repliement

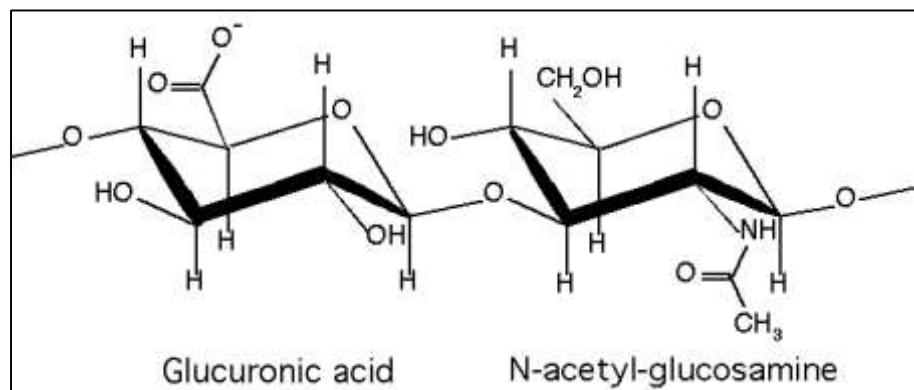


Figure 19 : Structure d'une unité disaccharidique d'acide hyaluronique

random coil) très hydrophiles rendant le milieu alentour viscoélastique. Cependant, les chaînes d'HA peuvent interagir faiblement entre elles et de manière transitoire *via* leurs côtés hydrophobes et/ou par des liaisons hydrogènes (Scott & Heatley 2002).

Les chaînes d'HA sont constituées de 200 à 20 000 répétitions de l'unité disaccharidique pour une masse moléculaire de 200 à 10^4 kDa. Dépliée, une chaîne peut mesurer de 2 à 25 μm (plus grand qu'une bactérie *E.coli* !). Le polymère a une certaine rigidité qui s'explique à la fois par la formation de réseaux transitoires de liaisons hydrogènes intramoléculaires au niveau des liaisons glycosidiques et également par la conformation « chaise » adoptée par l'acide glucuronique et la N-acétylglucosamine. Toutefois, cette conformation varie notamment lors d'interaction avec un partenaire protéique (Marković-Housley et al. 2000) au niveau de régions hautement flexibles et dynamiques (Day & Sheehan 2001).

c) Catabolisme

Le catabolisme des HA ainsi que des autres GAG est similaire puisqu'il fait intervenir des enzymes de dégradation possédant un spectre large de digestion. Ce processus sera détaillé pour les HA, KS, CS et HS/Hp dans une autre partie.

3. Les kératanes sulfates (KS)

La famille des kératanes sulfates (KS) aussi appelés kératosulfates a été découverte en 1939 par Suzuki dans la cornée de l'œil (Suzuki 1939). Ce n'est qu'en 1953 que Meyer et ses collaborateurs définissent la structure du polymère [$\rightarrow 3\text{Gal}\beta 1 \rightarrow 4\text{GlcNAc}\beta 1 \rightarrow$] (où Gal est un galactose et GlcNAc une *N*-acétylglucosamine) et lui donnent son nom de kératane sulfate (Meyer et al. 1953) (Figure 20). C'est le seul GAG à ne pas posséder d'acide uronique et donc

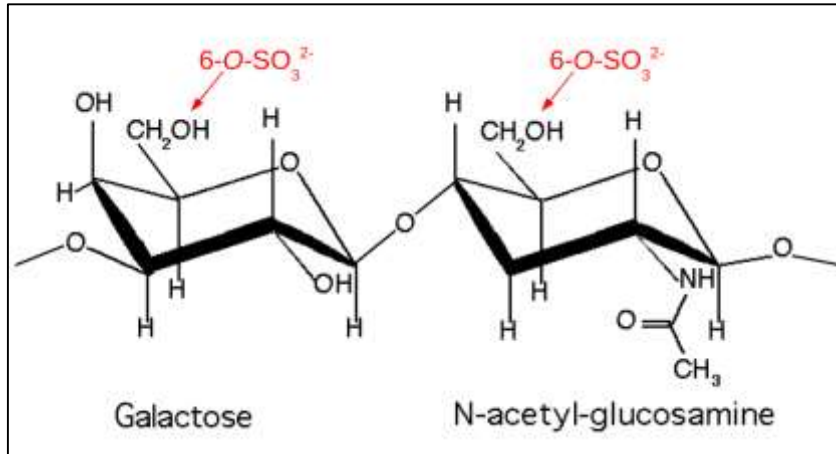


Figure 20 : Structure et modifications possible d'une unité disaccharidique de kératane sulfate

probablement pas de charge nette négative par disaccharide en solution.

Il existe 3 types de KS différenciés selon leur liaison à la protéine cœur, ils sont dénommés KS I, II ou III. Ils sont retrouvés dans des tissus différents, le KS I est principalement retrouvé dans la cornée de l'œil attaché aux

kératocan, lumican, mimécan principalement, et sa présence est essentielle dans la transparence de la cornée (Chakravarti et al. 1998) et de son hydratation. Les KS II sont retrouvés liés uniquement aux aggrécans et donc principalement dans les cartilages. Les KS III, quant à eux, sont retrouvés plutôt dans le cerveau au sein des PG phosphocan-KS, testican... Les KS I sont des glycans *N*-liés à une asparagine, les KS II sont liés à une galactosamine *N*-acétylée elle-même *O*-liée à une sérine/thréonine, et les KS III sont liés à une séquence Ser-*O*-Mannose.

a) Synthèse

(1) Initiation

L'initiation de la synthèse de chaîne de KS est différente selon leurs types. En effet, l'initiation de la polymérisation de la chaîne commence généralement avec l'ajout de sucre(s) liant la protéine cœur à la chaîne de KS, ce(s) sucre(s) est appelé (sont appelés) « linker(s) ». Ces différents *linkers* sont présentés (Figure 21) et peuvent être modifiés.

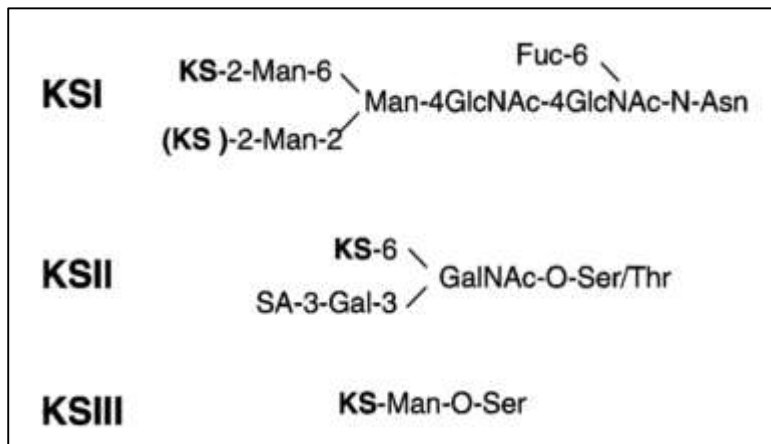


Figure 21 : Oligosaccharides de liaison des différents KS à leur protéine cœur. SA : acide sialique ; Fuc : fucose ; Man : mannose. D'après (Funderburgh 2002).

(2) Elongation

L'élongation des chaînes de KS consiste en l'ajout alternativement d'une *N*-acétylglucosamine et d'un galactose en $\beta 1 \rightarrow 4$ et $\beta 1 \rightarrow 3$ (Meyer et al. 1953). Cette élongation nécessite :

- une galactosyltransférase. Une enzyme extraite de cornée a été identifiée (Christner et al. 1979) et ressemble fortement à la β -1,4-galactosyltransférase ($\beta 4$ Gal-T1) retrouvée abondamment dans le sérum et le lait. Différentes isoformes existent avec différentes activités et seraient importantes dans l'élongation des différents KS (Ujita et al. 2000).

- une β -1,3-*N*-acétylglucosaminyltransférase ($\beta 3$ GlcNAcT). Cette enzyme n'est pas clairement identifiée, même si plusieurs d'entre elles ont pourtant été décrites dont une fonctionnant efficacement avec la $\beta 4$ Gal-T1. Cette $\beta 3$ GlcNAcT appelée iGnT a une expression tissulaire large, et une forte concentration en son transcrit ARN est retrouvée dans le cerveau, un organe dont la quantité en KS est importante (Fukuzumi et al. 2001).

Bien que des cibles responsables de la polymérisation des KS soient identifiées (Seko & Yamashita 2004), leur mode d'action n'est pas connu. Les enzymes agissent-elles de concert lors de l'élongation ? Existe-t-il d'autres combinaisons entre les enzymes des deux différentes classes ? Quel serait le lien avec l'expression en différents KS dans les différents tissus ?

(3) Terminaison

La terminaison de la polymérisation de la chaîne de KS est liée à ses différentes modifications et sera traité dans le point suivant.

b) Structures et modifications

Les KS sont des GAG particuliers car ils peuvent être biantennés dans le cas des KS I et II. Chez les KS I le premier mannose du côté de l'extrémité réductrice est mannosylé en position 3 et 6. C'est sur cette dernière position que l'élargissement de la chaîne est généralement faite alors qu'en position 3 la chaîne de KS est avortée après l'ajout d'un acide neuraminique à l'extrémité non-réductrice (Figure 22). Toutefois des chaînes de KS ont pu être mises en évidence en position 6 et 3 (Tai et al. 1996). Concernant les KS II, la ramification se fait au niveau de la *N*-acétylgalactosamine (GalNAc) *linker* où un galactose peut être adjoint en position 3, et une *N*-acétylglucosamine (GlcNAc) en position 6. Comme pour les KS, c'est à partir de la position 6 que se fait l'élargissement de la chaîne de KS. Le galactose en position 3 est également sialylé et cette modification est souvent retrouvée aux extrémités non-réductrices des polysaccharides (Funderburgh 2000) et pourrait agir comme un signal de terminaison de la polymérisation.

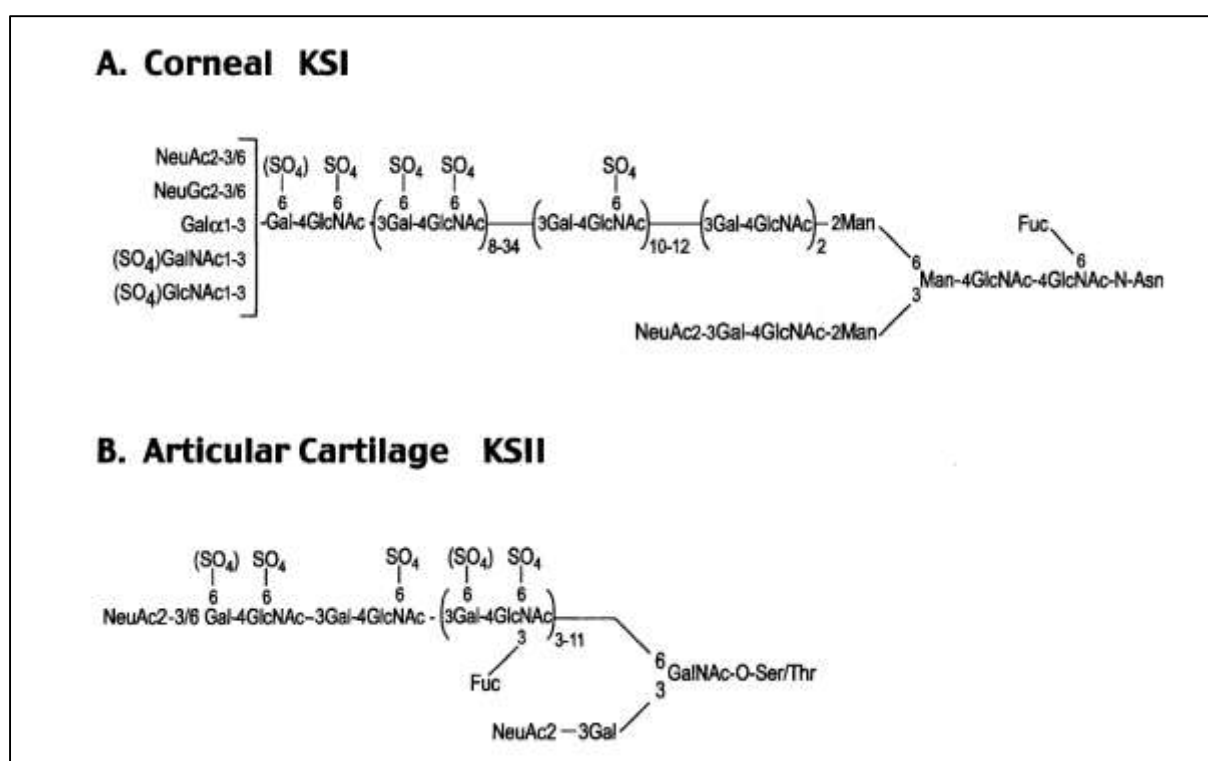


Figure 22 : Résumé des caractéristiques structurales des KS I (A) et II (B). NeuAc : acide neuraminique. D'après (Funderburgh 2000).

Les principales modifications des KS sont les sulfatations qu'ils peuvent subir à la fois sur le galactose et sur la GlcNAc en position 6. Les enzymes responsables de la 6-sulfatation des galactoses et des GlcNAc sont appelées KSGal6ST (pour kératane-sulfate-galactose-6-sulfotransférase) (Fukuta et al. 1997) et GlcNAc6ST ou Gn6ST (pour *N*-acétylglucosamine-6-*O*-sulfotransférase) respectivement. 5 gènes codant pour cette dernière enzyme ont été découverts et ont été caractérisés comme responsables des 6-sulfatations des GlcNAc aux extrémités non-réductrices dans des kératocytes (Nakazawa et al. 1998) (Yamamoto et al. 2001). La sulfatation en position 6 des GlcNAc est essentielle au cours de la biosynthèse, car une baisse de cette sulfatation provoque une baisse de la sulfatation des galactoses ainsi

qu'une réduction de la longueur de la chaîne de KS (Plaas et al. 2001). La structure finale des chaînes de KS (longueur, sulfatation, sialylation, fucosylation) a été montrée comme dépendante davantage du tissu dans lequel elles sont synthétisées que de la protéine cœur sur laquelle elles sont rattachées (Funderburgh 2002).

4. Les chondroïtines sulfates (CS), héparanes sulfates (HS) et Héparine (Hp)

a) Le tétrasaccharide linker

Le tétrasaccharide *linker* est une séquence conservée de 4 oses linéaires liée à une séquence consensus serylglycine (Muir 1958) de la protéine cœur. Cette séquence [...-GlcA β 1-3Gal β 1-3Gal β 1-4Xyl β 1-O-Ser] découverte en 1966 chez les CS (Rodén & Armand 1966) (Rodén & Smith 1966) est la même pour les CS, HS et Hp et a été conservée au cours de l'évolution (Yamada et al. 2002). Cette séquence va servir de substrat aux différentes enzymes de polymérisation des chaînes de CS ou de HS/Hp (Figure 23).

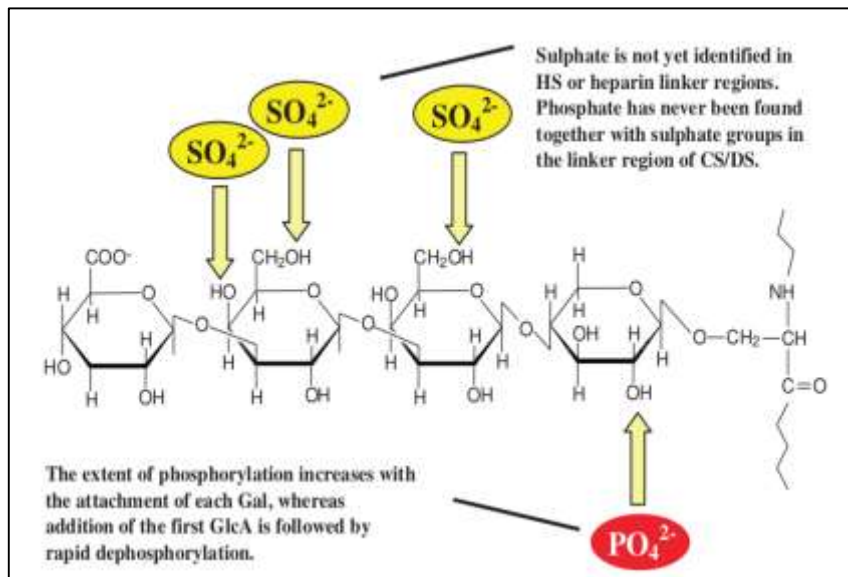


Figure 23 : Modifications éventuelles du tétrasaccharide linker. D'après (Prydz & Dalen 2000).

La formation de cette séquence de liaison est initiée par la xylosyltransférase I ou II (XylT-I ou -II) qui va catalyser la xylosylation de la sérine. La localisation de cette réaction est variable selon le type cellulaire mais a toujours lieu entre le reticulum endoplasmique (RE) et le réseau *cis*-golgien (Lohmander et al. 1989)

(Vertel et al. 1993). La suite de la polymérisation est l'adjonction séquentielle d'un premier puis d'un second galactose par les galactosyltransférases I et II respectivement (Okajima et al. 1999) (Bai et al. 2001), et s'effectue lors du transport du PG naissant au niveau des saccules de l'appareil de Golgi de la face *cis* à la région médiane (Prydz & Dalen 2000). Le dernier ajout osidique est effectué par la glucuronyltransférase I qui lie un glucuronate au dernier galactose du trisaccharide formé (Kitagawa 1998) lors de son acheminement de l'appareil de Golgi médian vers le *trans*-Golgi.

Le tétrasaccharide *linker* va pouvoir subir différentes modifications au cours de sa polymérisation. Une de ces modifications est la 2-O-phosphorylation du xylose indépendamment du GAG qui sera polymérisé (Fransson et al. 1985) (Moses et al. 1997). Dans le cas de la décorine, la phosphorylation du xylose est importante ($\approx 90\%$) jusqu'à la formation

du trisaccharide Gal-Gal-Xyl. Le xylose est ensuite déphosphorylé, et ce phénomène est concomitant avec l'ajout de l'acide glucuronique sur le dernier galactose (Moses et al. 1999). Cette phosphorylation transitoire pourrait être impliquée dans le transport intracellulaire et/ou dans le contrôle des modifications des chaînes glycaniques (Fransson et al. 2000). Les galactoses, quant à eux, peuvent être 4- et/ou 6-*O*-sulfatés, et ces sulfatations sont retrouvées exclusivement dans les chaînes de CS, même lorsque le PG est porteur des deux types de chaîne CS et HS (Ueno et al. 2001). Ces modifications des galactoses pourraient être des signaux de tri des CSPG ou des éléments contrôlant les glycosyltransférases impliquées dans la formation du tétrasaccharide et de la chaîne de GAG (Gulberti et al. 2005).

b) Les chondroïtines sulfates (CS)

Les CS ont été les premiers GAG découverts comme vu précédemment, et son motif disaccharidique [-4GlcA β 1-3GalNAc1 β -] fut déterminé en 1956 par Meyer *et al.* Il existe une grande hétérogénéité (de longueur, de modification) au sein des CS qui sont classés selon la sulfatation et l'épimérisation de l'unité disaccharidique qui les compose (Figure 24). Ils sont retrouvés abondamment chez les vertébrés, et plus largement dans le monde animal (*Caenorhabditis elegans*, *Drosophila melanogaster*...), notamment dans de nombreux organismes marins (requins, calmars, crabes, myxines, oursins, concombres de mer...) ainsi que chez certaines bactéries (*Escherichia coli* K4, *Pasteurella multocida*...).

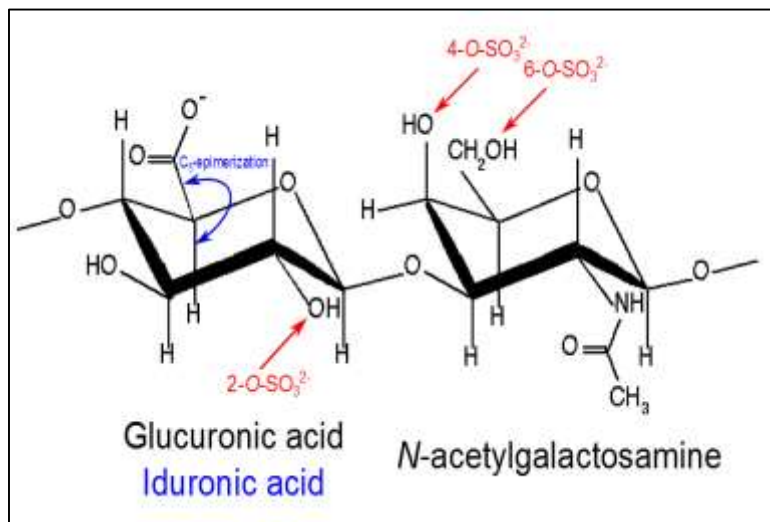


Figure 24 : Structure et modifications possibles d'une unité disaccharidique de chondroïtine sulfate

La composition des CS est fonction de l'organisme dans lequel ils se trouvent (Noulas et al. 2004), du tissu, du stade de développement (Bao et al. 2004), et de l'état physiopathologique de l'animal. Chez les mammifères, les CS sont quantitativement les GAG sulfatés les plus importants des matrices extra-cellulaires.

(1) Biosynthèse

La biosynthèse des CS débute par la liaison d'une première *N*-acétylgalactosamine (GalNAc) en β 1-4 au glucuronate du tétrasaccharide *linker* par la *N*-acétylgalactosaminyltransférase I ou II (GalNAcT I ou II) (Uyama et al. 2002) (Uyama et al. 2003) au niveau du réseau *trans*-golgien. La suite de la biosynthèse continue par l'ajout de manière séquentielle d'un acide glucuronique (GlcA) et d'une GalNAc en β 1-3 et β 1-4 respectivement

du côté non-réducteur. Ces réactions peuvent être catalysées par 6 enzymes différentes (Mikami & Kitagawa 2013). Parmi elles, sont regroupés les chondroïtines synthases (ChSy) -1, -2 et -3 et le facteur de polymérisation de chondroïtine (ChPF) possédant à la fois une activité GalNAcT-II et glucuronyltransférase-II (GlcAT-II). Enfin, deux chondroïtine-GalNAc-transférases (ChGn-1 et -2) complètent la famille, mais ces dernières ne sont capables de catalyser que le transfert de GalNAc sur la chaîne en cours d'élongation. Il a été montré que les ChSy et le ChPF pouvaient être coexprimés et former des complexes augmentant ainsi grandement leurs activités GlcAT-II et GalNAcT-II, la chaîne de CS finale étant dépendante des différentes combinaisons possibles entre les enzymes (Kitagawa et al. 2003) (Izumikawa et al. 2007) (Izumikawa et al. 2008) (Figure 25).

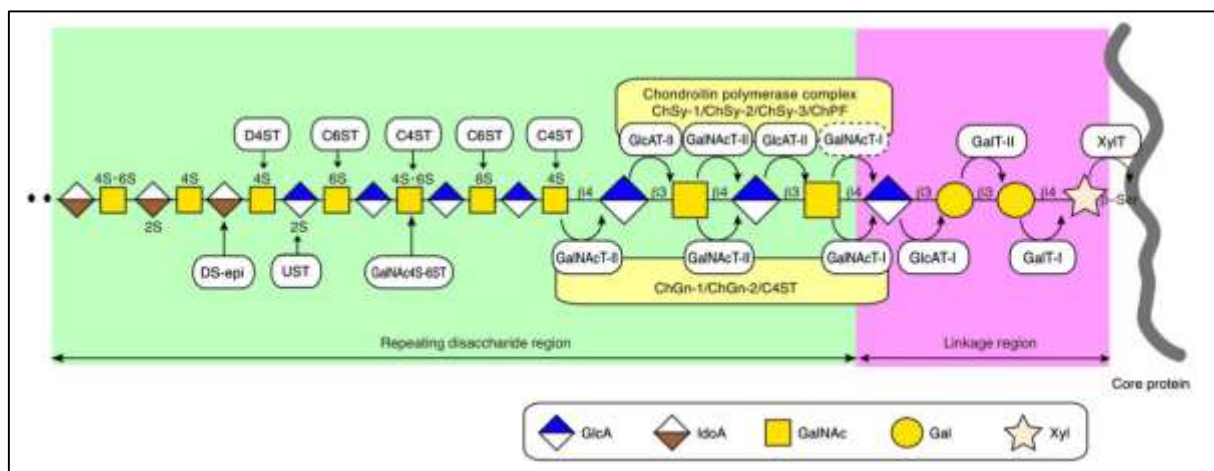


Figure 25 : Schéma conventionnel de la machinerie de la biosynthèse des CS. Voir texte pour abréviations. D'après (Mikami & Kitagawa 2013).

(2) Structures et modifications

Il existe à ce jour une dizaine de CS différents retrouvés dans les divers organismes étudiés. Chez l'Homme, 4 CS sont majoritairement retrouvés : les CS-A (ou chondroïtines 4-sulfate, C4S), les CS-C (ou chondroïtines 6-sulfates, C6S), les CS-B (ou dermatane sulfate, DS), et les CS-E (ou chondroïtines 4,6-disulfates). Ils se distinguent par le nombre de sulfates, leur position, ainsi que par l'épimérisation du carbone 5 de l'acide uronique (quand la liaison est équatoriale c'est un acide glucuronique GlcA, quand la liaison est axiale, c'est un acide iduronique IdoA). Quatre modifications sont importantes chez l'Homme :

-La C₅ épimérisation de l'acide uronique. Lorsqu'une chaîne de CS possède au moins un acide iduronique, c'est à dire que l'acide glucuronique ajouté lors de l'étape d'élongation a subi une épimérisation, cette chaîne est appelée dermatane sulfate (DS). Il existe 2 isoformes de l'enzyme responsable de cette épimérisation nommées dermatane-sulfate-épimérase-1 et -2 (DS-épi1, DS-épi2) (Pacheco et al. 2009). La quantité d'IdoA dans une chaîne de DS peut fortement varier selon le tissu à partir duquel elle est extraite et peut représenter de 7 à 93 % du total en acide uronique. La présence d'iduronates chez les DS rend la chaîne plus flexible par rapport aux chaînes non-épimérisées. En effet, alors que les GlcA

(comme les GalNAc) adoptent une conformation spatiale 4C_1 (dite de type « chaise ») (Casu et al. 1986) (Ferro et al. 1990), les IdoA, eux, peuvent en plus adopter une autre conformation « chaise » 1C_4 et une conformation 2S_0 (*skew boat*). Les IdoA en solution ne vont pas adopter une seule conformation fixe, le phénomène est dynamique et un IdoA peut passer par différentes structures tridimensionnelles. La structure même de ces IdoA est à la base de certaines interactions DS-protéines (Trowbridge & Gallo 2002).

-La 4-O-sulfatation des GalNAc. Les chaines possédant ces sulfatations sont de type CS-A et ce sont les modifications des CS les plus abondantes. 4 enzymes sont responsables de la 4-O-sulfatation des GalNAc à partir du PAPS qui est le donneur cellulaire de sulfates. Parmi ces 4 enzymes, on retrouve 3 chondroïtine-4-sulfotransférases (C4ST-1, -2 et -3) spécifiques des CS et 1 dermatane-4-sulfotransférase (D4ST) spécifique des DS. La C4ST-1 est particulièrement importante car lors de mutation ou de *knock out* du gène codant l'enzyme chez la souris, les autres isoformes C4ST-2 et -3 ne compensent pas l'absence d'activité de la C4ST-1. De plus, la C4ST-1 possède un rôle plus large puisque non seulement la 4-O-sulfatation des chaines diminue, mais la quantité de chaine de CS produite est également plus faible dans tout l'organisme (Kluppel et al. 2005). Enfin, la D4ST contrairement aux C4ST est capable de catalyser la 4-O-sulfatation des GalNAc lorsqu'un IdoA est adjacent. Cette sulfatation empêche l'épimérase de re-épimériser l'IdoA en GlcA et bloque donc l'acide uronique en son épimère iduronique (Mikami et al. 2003).

-La 6-O-sulfatation des GalNAc. Cette modification des chaines les classe dans la famille des CS-C. 3 enzymes sont responsables de cette réaction. Les chondroïtine-6-sulfotransférases (C6ST-1 et -2) transfèrent un sulfate du PAPS vers le groupement porté par le carbone 6 de la GalNAc. La C6ST-1 a montré un spectre de réaction plus large car elle est en effet capable de 6-O-sulfater les galactoses dans les chaines de KS (Uchimura et al. 1998). Chez l'Homme aucune 6-O-sulfotransférase spécifique des DS n'a été décrite, bien que la C6ST-1 recombinante ait montré une activité 6ST envers les DS *in vitro* (Yamada et al. 2004). Enfin, une dernière C6ST spécifique des GalNAc4S a été caractérisée et est appelée *N*-acétylgalactosamine-4-sulfate-6-sulfotransférase (GalNAc4S-6ST). Elle est responsable de la 6-O-sulfatation de la GalNAc4S, et lorsque ceux-ci sont retrouvés abondamment au sein de la chaine de CS, cette chaine est décrite comme CS-E. Ces chaines sont exclusivement retrouvées attachées à la serglycine dans les mastocytes des muqueuses. Toutefois, son expression est très large au sein de l'organisme car les GalNAc4S,6S sont retrouvées en abondance en bout de chaine du côté non-réducteur. La présence de cette disulfatation particulière est par exemple 60 fois plus abondante à l'extrémité non-réductrice de la chaine de CS des agrégans qu'au milieu (Midura et al. 1995). Il constituerait un signal de terminaison pour les enzymes impliquées dans l'élongation de la chaine de CS. Cette hypothèse est renforcée par les travaux de Ohtake-Niimi et ses collaborateurs qui ont montré chez les souris *KO* pour le gène codant pour la GalNAc4S-6ST qu'elles possédaient des chaines de CS plus longues que les souris non-mutées (Ohtake-Niimi et al. 2010).

-La 2-O-sulfatation des acides uroniques. Une seule enzyme a été caractérisée comme responsable de la 2-O-sulfatation des acides glucuroniques et iduroniques (Kobayashi

et al. 1999). Cette uronyl-2-*O*-sulfotransférase (UST) a une meilleure activité pour les DS et lorsqu'une GalNAc4S est adjacente à l'acide uronique. Ce serait la dernière enzyme à intervenir au niveau intracellulaire dans l'élaboration de la chaîne finale de CS/DS puisqu'elle agit préférentiellement sur les unités disaccharidiques épimérisées et/ou sulfatées, vraisemblablement tardivement au niveau du réseau *trans*-golgien (Figure 26).

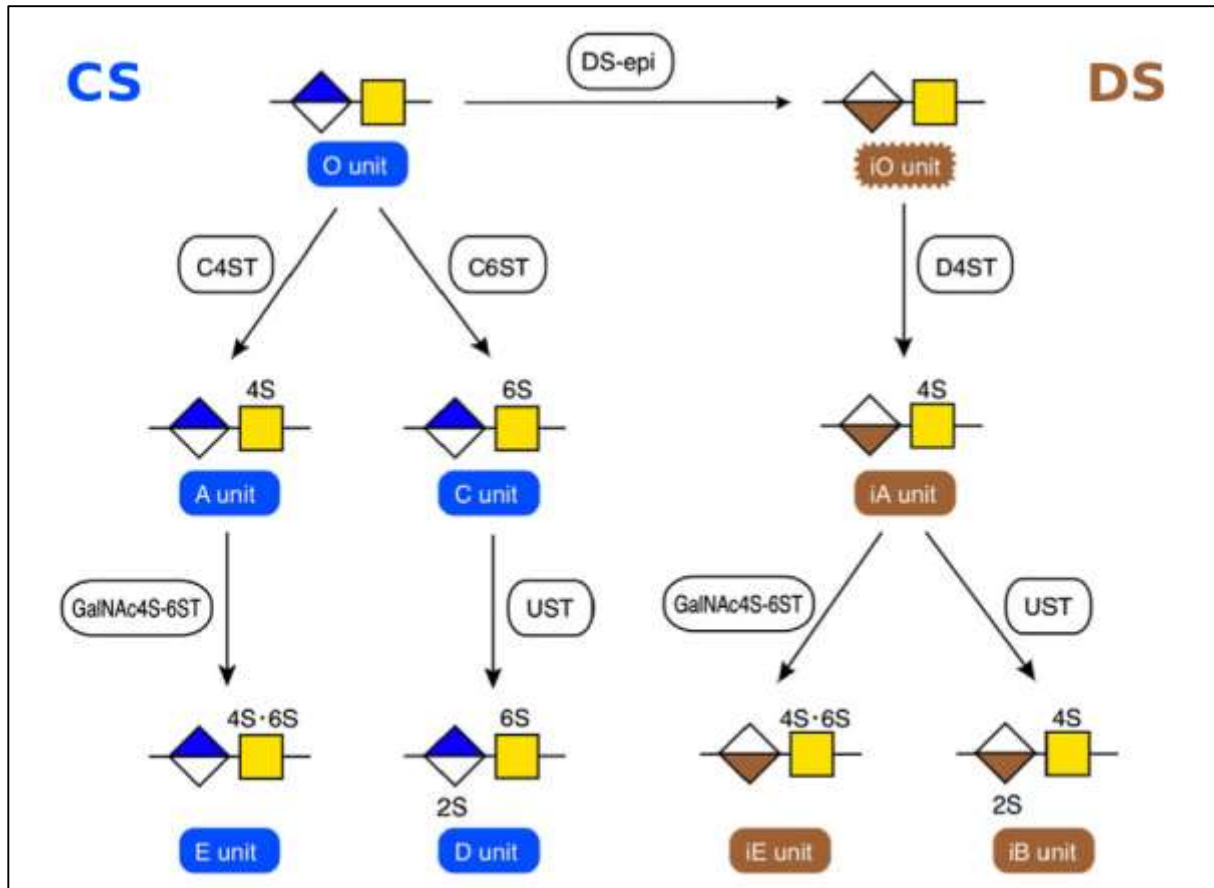


Figure 26 : Diagramme schématisant des voies de modification biosynthétiques des chaînes de CD/DS. D'après (Mikami & Kitagawa 2013).

D'autres modifications absentes chez les mammifères comme des 3-*O*-sulfatations, fucosylations, *N*-sulfatations... sont retrouvées chez différents invertébrés marins mais ne seront pas détaillées dans cette partie.

c) Les héparanes sulfates (HS) et l'héparine (Hp)

L'héparine fut découverte par McLean en 1916 dans le foie (du grec *hepar* signifiant foie) de chien (pour histoire voir (Marcum 2000)) grâce à ses propriétés anticoagulantes. Ce n'est qu'en 1935 qu'elle fut classée comme appartenant à la famille des GAG. Les héparanes sulfates, quant à eux, n'ont été découverts qu'en 1948 en tant que sous-produit de la purification d'héparine. La découverte fortuite de ces derniers peut s'expliquer en partie parce qu'héparine et héparanes sulfates sont très proches. En effet, ils possèdent un squelette osidique identique de type $[-4\text{GlcA}\beta 1-4\text{GlcNAc}\alpha 1-]$, et sont modifiés par les mêmes enzymes (Figure 27). La différence entre ces deux polymères vient du fait qu'ils ne sont pas exprimés

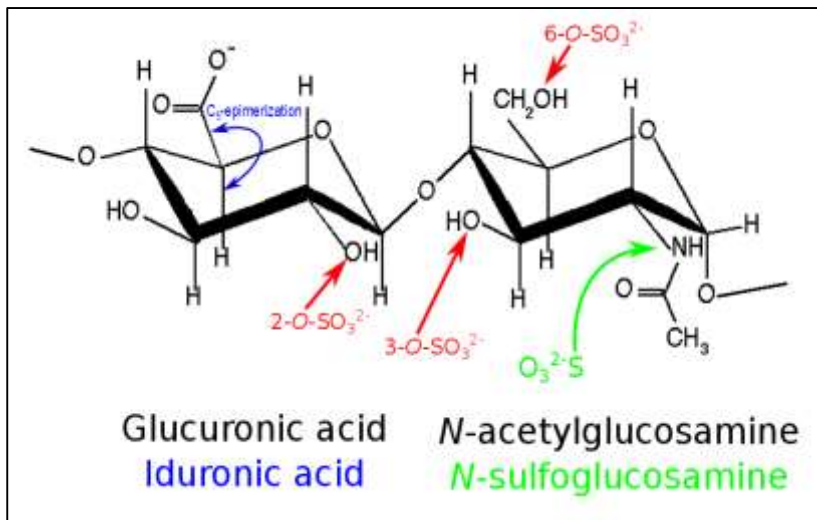


Figure 27 : Structure et modifications possibles d'une unité disaccharidique d'héparane sulfate/héparine

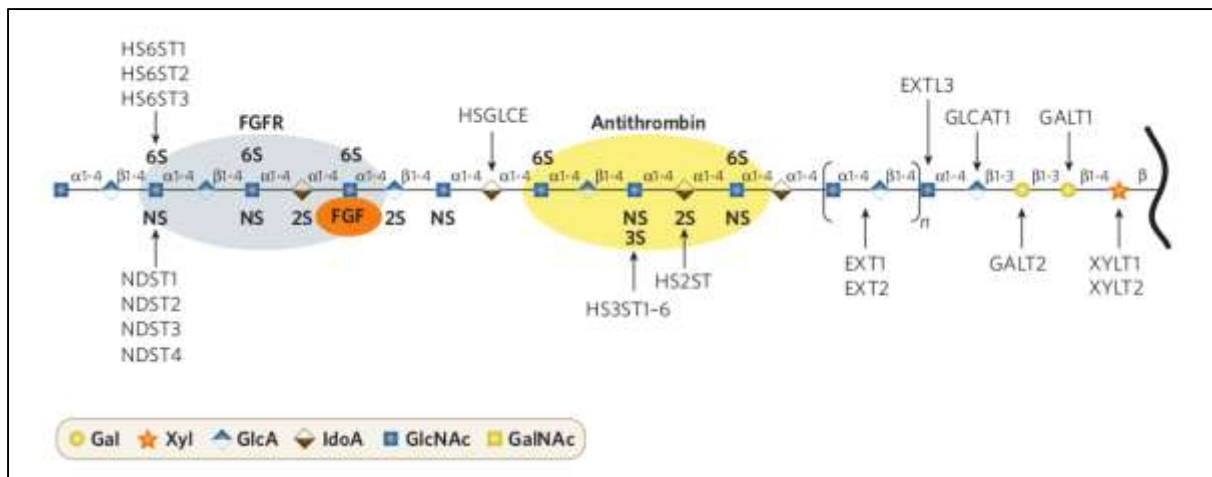
dans les mêmes tissus. Si les héparanes sulfates sont exprimés de manière ubiquitaire, les chaînes d'héparine, quant à elles, sont exprimées exclusivement dans les mastocytes des tissus conjonctifs. L'expression des enzymes impliquées dans la biosynthèse des Hp est différente des autres cellules de l'organisme,

expliquant la différence de structure finale entre les 2 polymères.

Les HS/Hp sont les GAG les plus modifiés comme nous le verrons ce qui leur permet d'interagir avec une grande diversité de protéines. De plus, les HS sont principalement retrouvés attachés au niveau de la membrane plasmique, en contact direct avec la cellule et son environnement, une place stratégique qui explique ses nombreux rôles physiologiques.

(1) Biosynthèse

L'initiation de la chaîne de HS/Hp commence par l'ajout d'une première *N*-acetylglucosamine (GlcNAc) au tétrasaccharide *linker* par une enzyme transmembranaire avec une activité *N*-acetylglucosaminyltransférase-I (GlcNAcT-I). Deux enzymes de la famille des « exostosine-like » (EXTL) sont capables de catalyser cette réaction : EXTL-2 et EXTL-3 (Figure 28). Après l'ajout de cette GlcNAc sur le tétrasaccharide *linker*, c'est un complexe de 2 enzymes membranaires qui va être responsable en grande partie de l'élongation de la chaîne HS/Hp. Ce complexe des deux enzymes exostosine-1 et -2 (EXT1 et EXT2) possède l'activité GlcAT-II et GlcNAcT-II nécessaire à la polymérisation des HS/Hp. EXTL-1 et -3 ont également montré une activité GlcNAcT-II et peuvent jouer un rôle dans l'élongation des chaînes. L'étude de souris *KO* pour les gènes codant pour les EXTL-2 et -3, impliquées dans l'initiation des chaînes, a été montrée *in vitro* que l'élongation pouvait tout de même être assurée par le complexe EXT1/EXT2 (Kim et al. 2003). Lors de mutation de EXTL-3, les chaînes d'HS caractérisées étaient plus longues (Busse et al. 2007), ce qui signifie que l'enzyme jouerait un rôle dans la régulation de la taille des chaînes (Nadanaka & Kitagawa 2008). Toutefois, les mécanismes de terminaison de synthèse des HS et Hp sont largement méconnus.



(2) Structures et modifications

-La N-déacétylation/N-sulfatation des GlcNAc par des *N*-déacétylases/*N*-sulfotransférases (NDST). 4 isozymes bifonctionnelles existent et catalysent la réaction de *N*-déacétylation puis de *N*-sulfotransfert sur les glucosamines. Leur expression dépend de leur localisation ainsi que du stade de développement. En effet, alors que les transcrits codant les NDST-1 et -2 sont retrouvés partout dans les tissus embryonnaires et adultes avec différents niveaux d'expression (Kusche-Gullberg 1998), ceux des NDST-3 et -4 sont exprimés principalement au niveau de l'embryon et du cerveau (Aikawa & Esko 1999) (Aikawa et al. 2001). La génération de souris *KO* aux NDST a permis de mieux cerner leurs différents rôles. La NDST-1 est essentielle à la survie de l'animal qui meurt à l'état embryonnaire ou à la naissance en absence de NDST-1 (Ringvall et al. 2000) (Fan et al. 2000). Les souris *ndst-2*^{-/-} quant à elles sont viables mais des défauts dans les chaînes d'héparine sont observés. Ces dernières ne retiennent plus toutes les enzymes et protéines dans les granules de sécrétion des mastocytes (des tissus conjonctifs) ce qui engendre des troubles immunitaires, et notamment dans les phénomènes inflammatoires. La NDST-2 serait donc essentielle dans la formation des chaînes d'héparine. De manière intéressante, la NDST-2 est surtout retrouvée à l'état de transcrit ARNm dans la plupart des tissus de l'organisme et non au niveau protéique. Il existerait une régulation en dehors des mastocytes plutôt traductionnelle que transcriptionnelle qui expliquerait également l'absence de phénomène compensatoire dans les souris *ndst-1*^{-/-} (Grobe et al. 2002). Selon l'isoforme de l'enzyme, les activités *N*-déacétylases et *N*-sulfotransférases sont différentes. Dans le cas de la NDST-3 l'activité *N*-déacétylase est plus importante que l'activité *N*-sulfotransférase, ce qui peut expliquer la

présence de glucosamines non-substituées (GlcNH_3^+) impliquées dans des phénomènes d'interaction (Norgard-Sumnicht & Varki 1995) (Vanpouille et al. 2007). La NDST-2 a une activité *N*-déacétylase et *N*-sulfotransférase plus importante que la NDST-1, ce qui explique la plus forte quantité de glucosamines *N*-sulfatées (GlcNS) dans les chaînes Hp par rapport aux HS. Les NDST forment des séquences contiguës de GlcNS qui vont être des substrats préférentiels pour les enzymes de modification plus en aval. Ce sont elles qui forment l'ébauche des chaînes d'HS/Hp.

-La C_5 épimérisation des acides glucuroniques (GlcA) en acides iduroniques (IdoA) par la C_5 -épimérase (HSepi ou HSGLCE). Cette épimérase est différente de celle responsable de l'épimérisation des GlcA en IdoA des DS. Une seule isoforme de l'enzyme a été mise en évidence (Lindahl & Bäckström 1972) chez les mammifères et catalyse la réaction d'épimérisation des GlcA en IdoA, et inversement *in vitro*. L'HSepi reconnaît comme substrat des GlcA avec au moins une glucosamine *N*-sulfatée adjacente du côté non-réducteur (Jacobsson et al. 1984). Le pourcentage en IdoA varie selon le type de chaîne et le tissu dans lequel il est retrouvé. Généralement, 30 à 55% d'IdoA sont retrouvés dans les chaînes d'HS, et de 50 à 90% dans les chaînes d'Hp (Taylor et al. 1973). Ce fort taux dans les chaînes d'héparine vient du fait que le polysaccharide est *N*-sulfaté par la NDST-2, plus efficace que la NDST-1, offrant ainsi plus de séquences substrats pour l'HSepi. Cette dernière est essentielle dans le développement puisque les souris *KO* pour le gène codant l'HSepi meurent à la naissance. Elles présentent de nombreuses anomalies au niveau de l'iris, du squelette, les poumons sont immatures et les reins ne sont pas développés (Li 2010). Les cellules de ces souris sont remis en culture afin d'étudier leurs chaînes d'HS et d'Hp. Les deux types de GAG après analyse montrent une structure similaire, avec une forte *N*-sulfatation et 6-*O*-sulfatation des glucosamines (Feyerabend et al. 2006), suggérant un phénomène de « compensation » de l'absence en IdoA et en 2-*O*-sulfatation qui est sa modification principale. Comme dit précédemment dans la partie des dermatanes sulfates, les IdoA ont une flexibilité conformationnelle plus importante que les autres oses de la chaîne de GAG. Cette conformation de l'IdoA est également fondamentale pour l'interaction des HS/Hp avec un partenaire protéique (Jia et al. 2009).

-La 2-*O*-sulfatation par la 2-*O*-sulfotransférase (HS2ST). Une seule isoforme de l'enzyme existe chez les mammifères. La 2-*O*-sulfatation a lieu principalement au niveau des IdoA dans les chaînes d'HS/Hp et la réaction empêche l'épimérisation retour en acide glucuronique car les acides iduroniques 2-*O*-sulfatés (IdoA2S) ne sont pas des substrats pour l'HSepi. Les 2 enzymes forment un complexe favorisant ainsi la 2-*O*-sulfatation des IdoA (Pinhal et al. 2001) (Esko & Selleck 2002). Lors de *KO* pour le gène codant la HS2ST chez la souris, la progéniture meurt au stade néonatal. Les souriceaux présentent une agénèse rénale, des malformations squelettiques, de l'iris... Ces défauts sont également retrouvés chez les souris *hs2st*^{-/-} qui, du fait de l'absence d'IdoA, ne possèdent quasiment pas de 2-*O*-sulfatation. Cette sulfatation est impliquée et nécessaire dans l'interaction et la signalisation de plusieurs molécules comme le FGF-2, le GDNF (facteur neurotrophe dérivé de la glie impliqué dans le développement rénal) (Li 2010)... Les chaînes d'HS/Hp issues des cellules *hs2st*^{-/-} montrent une

forte augmentation de leur taux de *N*-sulfatation et 6-*O*-sulfatation comme pour « compenser » au niveau charge électrique l'absence de 2-*O*-sulfatation. La HS2ST est également capable de 2-*O*-sulfater les GlcA, son activité envers cet acide uronique est largement minoritaire dans des conditions « physiologiques » (Rong et al. 2001). Ces GlcA2S sont principalement retrouvées au niveau du cerveau (Lindahl et al. 1995), mais la régulation de ce phénomène est encore mal connue à ce jour.

-La 6-*O*-sulfatation par les 6-*O*-sulfotransférases (HS6ST). 4 isoformes (HS6ST-1, -2, -2S, -3) sont responsables de l'ajout d'un groupement sulfate en position 6 d'une glucosamine à partir du PAPS (Habuchi et al. 2000) (Habuchi et al. 2003). L'expression de chaque isoforme varie selon le tissu et le stade de développement (Sedita et al. 2004). Leurs affinités en fonction du substrat varient également d'une isoforme à une autre. En effet, la HS6ST-1 est capable de 6-*O*-sulfater les GlcNS lorsqu'un IdoA est adjacent à cette glucosamine. En revanche, les HS6ST-2 et -3 peuvent catalyser la réaction en absence de cet épimère. Cependant, pour que cette réaction ait lieu, la HS6ST-2 nécessite une forte concentration en disaccharide GlcA-GlcNS contrairement à la HS6ST-3. La présence de ce groupement sulfate en position 6 de la glucosamine va conférer à la chaîne d'HS/Hp des propriétés d'interaction et d'activation de molécules signal impliquées dans des phénomènes de migration, prolifération, adhésion (Pye et al. 1998) (Pye et al. 2000) (Mahalingam et al. 2007)... Si la réaction de 6-*O*-sulfatation a lieu préférentiellement sur les glucosamines *N*-sulfatées, il a été mis en évidence que la HS6ST-1 avait également pour substrat la GlcNAc (Zhang et al. 2001). La formation de ces GlcNAc6S est essentielle puisqu'elles jouent des rôles physiologiques fondamentaux de par leurs interactions avec l'antithrombine (rôle dans la coagulation), certains facteurs de croissance *etc.* Les souris *hs6st*^{-/-} ne sont pas viables et ont une mort embryonnaire ou néonatale (Habuchi et al. 2007).

-La 3-*O*-sulfatation par les 3-*O*-sulfotransférases (HS3ST). Elles forment une famille de 7 membres (HS3ST-1, -2, -3A, -3B, -4, -5, -6) (Shworak et al. 1999) (Xu et al. 2005) dont chacune a un niveau d'expression différent selon le tissu dans lequel elles sont exprimées et le stade de développement (Kusche-Gullberg & Kjellén 2003). Les HS3ST sont responsables de l'ajout de 3-*O*-sulfates au niveau des glucosamines, chacune d'entre elle ayant une activité différente selon le substrat. Cette réaction est la plus rare parmi les différentes modifications vues précédemment, mais aussi l'une des plus régulées (impliquant 7 isozymes différentes). Cela permet de 3-*O*-sulfater des séquences uniques au niveau de la chaîne d'HS/Hp dans des tissus bien particuliers. Les chaînes ainsi modifiées vont ainsi posséder un motif d'interaction critique pour des protéines fonctionnelles (par exemple avec l'antithrombine-III qui reconnaît un pentasaccharide particulier possédant une GlcNS,6S,3S dans la chaîne d'héparine (Petitou et al. 2003) : rôle dans l'hémostase). Certains virus également nécessitent la présence de 3-*O*-sulfate dans la chaîne d'HS pour s'adsorber au niveau de la membrane plasmique, c'est le cas par exemple du virus de l'herpès (HSV-I) (Shukla et al. 1999).

-La notion de « gagosome ». La synthèse (polymérisation et modifications) des chaînes de GAG, et surtout des HS et Hp qui sont les plus étudiés, est un processus rapide, efficace et étrangement reproductible pour un processus non-codé génétiquement. En effet, les

caractéristiques des HS sont généralement les mêmes dans un même tissu. Le gagosome serait un macrocomplexe regroupant les enzymes de synthèse et serait à l'origine de cette reproductibilité dans la polymérisation et modification des chaînes d'HS. Les arguments en faveur de ce concept sont les interactions observées entre certaines enzymes comme EXT-1, EXT-2 et les NDST (McCormick et al. 2000) (Kobayashi et al. 2000) (Presto et al. 2008), et entre la HSEpi et la HS2ST (Pinhal et al. 2001). Des modulateurs encore non décrits à ce jour pourraient être responsables du maintien cohérent de ce macrocomplexe. L'expression des différentes enzymes et la stœchiométrie au sein du supposé complexe joueraient un rôle clé dans l'établissement de structures reproductibles (Esko & Selleck 2002) (Figure 29).

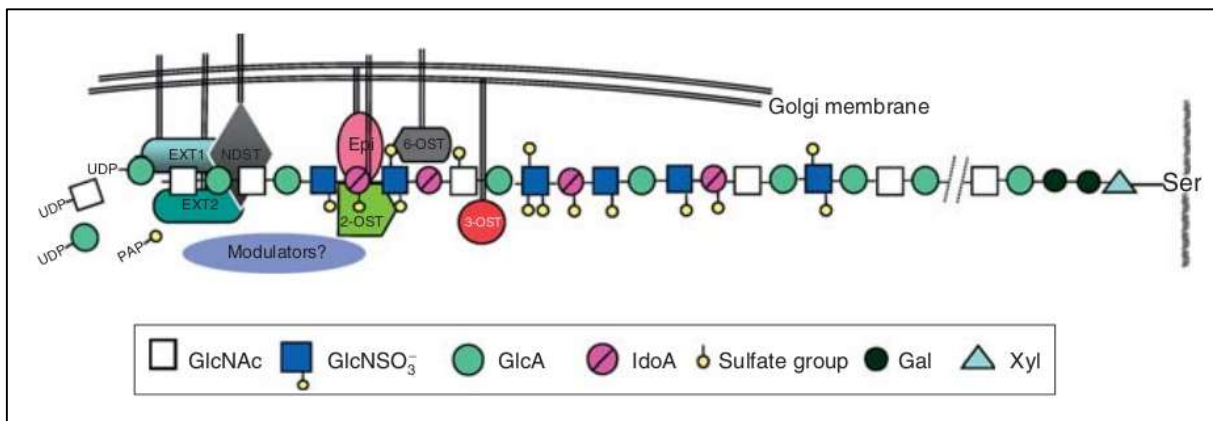


Figure 29 : Modèle théorique du gagosome pour la synthèse d'HS. D'après (Lindahl & Li 2009).

Les chaînes d'HS et d'Hp en fin de biosynthèse reflètent l'état physiologique du tissu dans lequel elles sont exprimées. Leurs caractéristiques (longueur, sulfatation, épimérisation) dépendent de l'expression des différentes isoformes des enzymes impliquées dans cette voie de biosynthèse. Les modifications des chaînes ont lieu majoritairement dans des domaines présentant des GlcNS contigus (les domaines *NS*) espacés par des domaines non-modifiés (les domaines *NA*). Aux frontières de ces domaines, la *N*-sulfatation des glucosamines est moins régulière et forment des domaines de transition (domaines *NS/NA*). Ce sont donc les NDST qui forment un premier plan de modification des chaînes HS/Hp. Cependant, toutes les modifications des chaînes ne sont pas exclusivement intracellulaires. De nouvelles enzymes appelées « Sulfs » (pour sulfatases) ont été mises en évidence au niveau extracellulaire et modifient les chaînes d'HS exprimées sur les surfaces membranaires :

-La famille des Sulfs a été découverte d'abord chez la caille en 2001 (Dhoot et al. 2001) puis chez l'Homme en 2002 (Morimoto-Tomita et al. 2002). Deux isoformes existent (HSulf-1 et HSulf-2 chez l'humain) et possèdent une activité 6-*O*-endosulfatase, c'est-à-dire qu'elles sont capables d'enlever sélectivement des 6-*O*-sulfates des glucosamines au sein des domaines *NS* (Staples et al. 2011). L'expression de ces enzymes sécrétées a un rôle important car elles peuvent moduler les interactions avec différents partenaires requérant la présence de groupements 6-*O*-sulfate (Seffouh et al. 2013). La mort néonatale des souris *KO* pour les gènes codant les 2 isoformes de Sulf montre également un rôle important lors de l'embryogenèse (Holst et al. 2007). La régulation de l'expression des Sulfs est un domaine

d'investigation important qui a déjà montré le rôle de celles-ci dans des cancers lors de dérégulation (Khurana et al. 2013).

La diversité structurale des chaînes d'HS/Hp prend sa source au cours de leur biosynthèse, mais la régulation de cette diversité est également importante après leur polymérisation au niveau des surfaces cellulaires et des MEC.

5. Le catabolisme des GAG *in vivo*

Les GAG sont retrouvés de manière ubiquitaire dans l'organisme et différentes voies de dégradation existent afin de pouvoir tous les renouveler quel que soit le tissu. Une première étape est sa dégradation *in situ* dans les MEC et dans le liquide synovial où il est ensuite drainé dans le système lymphatique (Antonias et al. 1973). Les chaînes polysaccharidiques sont catabolisées en de plus petits fragments qui vont ensuite passer dans le flux sanguin (Laurent & Laurent 1981) pour être captés soit par le foie (80%) où la dégradation y sera totale, soit par les reins où ils vont ensuite être excrétés par voie urinaire. Des fractions de GAG peuvent être internalisées et dégradées à l'intérieur du tissu, c'est le cas au niveau de la peau et des articulations. Cette internalisation peut être médiée par des récepteurs comme le CD44 interagissant avec l'HA (Lesley et al. 2000) ou par pinocytose (endocytose constitutive).

Les chaînes sont internalisées dans des endosomes qui vont fusionner avec les lysosomes contenant les enzymes de dégradation des GAG. Cette dégradation est assurée par des hydrolases acides spécifiques de la stéréoconfiguration des sucres, de la liaison glycosydique, des différents motifs de modification, de l'ose voisin... Le processus de dégradation est ordonné : dans un premier temps, ce sont des endoglycosidases spécifiques qui vont agir sur les chaînes. Les héparanases vont digérer les HS/Hp en oligosaccharides, et les différentes hyaluronidases et hexosaminidases vont digérer les HA, CS/DS et KS (Stern 2003) (Kaneiwa et al. 2010) (Ludolph et al. 1981). Après l'action des endoglycosidases au niveau du lysosome, ce sont des exoglycosidases qui vont pouvoir catalyser les réactions de désulfatations, réacétylation et dégradation des oligosaccharides en monosaccharides qui serviront dans l'anabolisme de nouvelles chaînes une fois activés par l'ajout d'UDP. Les exoglycosidases agissent au niveau de l'extrémité non-réductrice de l'oligosaccharide, et l'action préliminaire des endoglycosidases permet d'augmenter la quantité de substrats au sein des lysosomes. Les sulfatases responsables de la désulfatation des oligosaccharides peuvent avoir plusieurs substrats (cas de l'iduronate-2-sulfatase agissant sur les HS/Hp et DS, de la GalNAc-6-sulfatase ayant pour substrat les CS-C et les KS...). Après la désulfatation du résidu en position non-réductrice, la liaison avec l'ose voisin est rompue par une β -glucuronidase, α -iduronidase, β -galactosidase, β -N-acétylhexosaminidase, ou α -N-acétylhexosaminidase selon la nature de l'ose et de la liaison osidique. Des mutations dans ces enzymes de dégradation sont responsables de pathologies appelées mucopolysaccharidoses (MPS) caractérisées par l'accumulation des GAG (Figure 30).

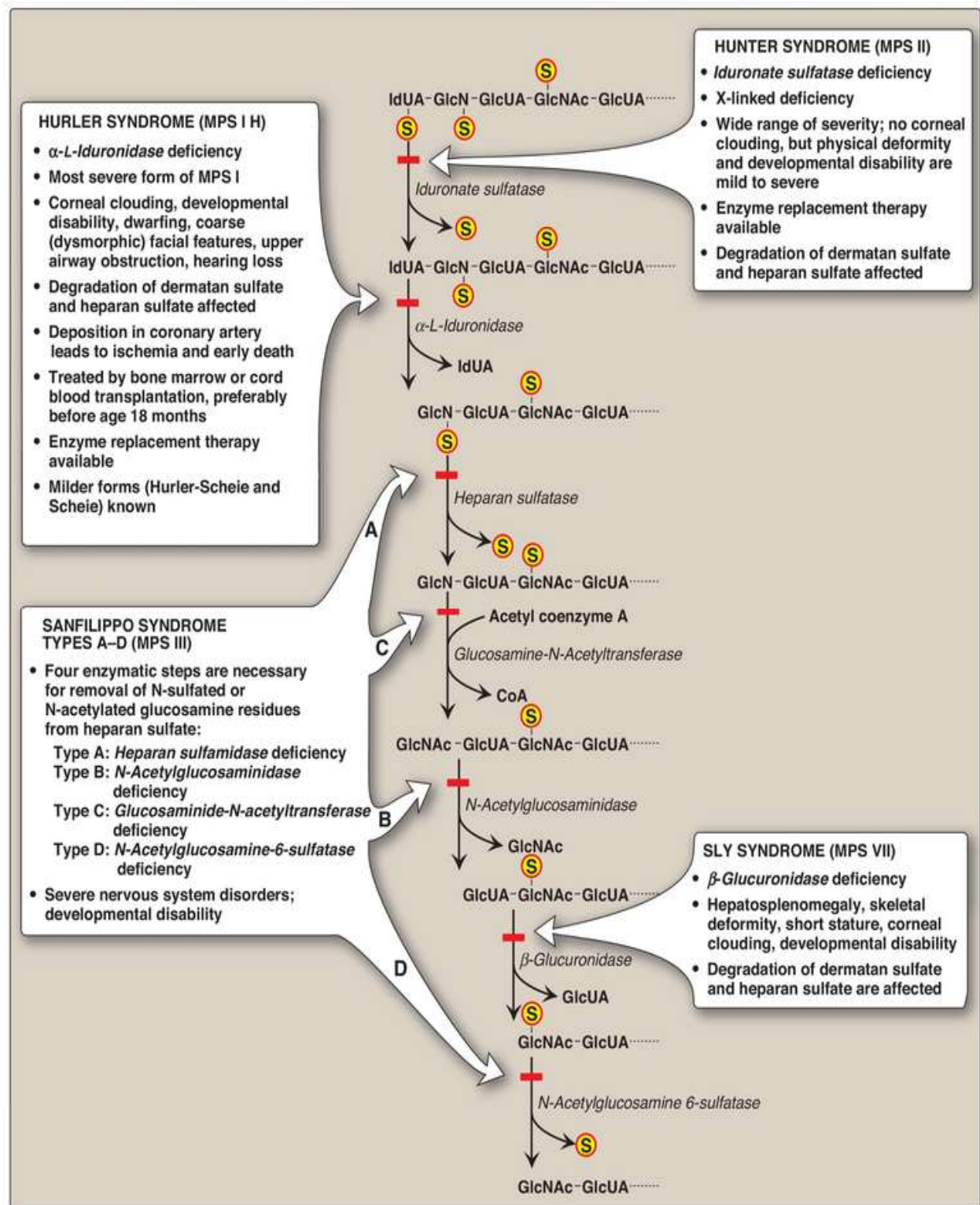


Figure 30 : Etapes de la dégradation des HS in cellulo. Enzymes impliquées dans la dépolymérisation et pathologies associées à leur déficience. D'après Biochemistry 6th edition.

Certaines des endoglycosidases comme l'héparanase peuvent être sécrétées et être en contact direct avec la MEC. Ces-dernières possèdent également une activité enzymatique à pH neutre, et elles sont impliquées dans la déstabilisation de certaines MEC pour des phénomènes de migration cellulaire, et de fertilisation (Lin et al. 1994) par exemple. Leur action peut permettre de relarguer en solution des cytokines retenues par les constituants de

la MEC et jouer des rôles dans des proliférations contrôlées (Gingis-Velitski et al. 2004) ou dérégulées comme c'est le cas dans certains cancers (Vlodavsky et al. 1999).

Les GAG peuvent également être dégradés de manière non-enzymatique. En effet, dans des processus inflammatoires (comme l'arthrite) ou lors d'autres maladies, des espèces réactives de l'oxygène (ROS) peuvent se former et cliver les GAG, déstabilisant ainsi les cartilages ou tissus impliqués lors de cette inflammation (Al-Assaf et al. 2000). Les fragments de GAG ainsi formés vont se détacher de leur environnement et être recaptés par les cellules ou passer dans la circulation lymphatique. Les leucocytes sont également capables de générer des espèces chimiques réactives envers les GAG aboutissant à leur dégradation (Rees et al. 2007).

D. Rôles des PG

Le rôle des PG est extrêmement varié dans l'organisme lors du développement mais également tout au long de la vie de l'individu et nous détaillerons son implication dans les phénomènes immunitaires et principalement lors de l'inflammation et du guidage leucocytaire. Enfin, les PG sont aussi impliqués dans de nombreux événements pathologiques dont nous verrons quelques exemples. Ces rôles peuvent être liés à la protéine cœur ou à leur chaîne de GAG.

1. Rôles de la protéine cœur...

De par leur grande diversité et leur localisation ubiquitaire, les protéines cœurs des PG jouent des rôles dans de très nombreux processus physiologiques. Nous ne pourrions tous les détailler ici, mais certains de ces rôles seront décrits dans cette partie.

a) ... dans le choix de la chaîne de GAG

L'un des rôles fondamentaux des protéines cœur est la participation de celles-ci dans le choix du GAG qui va être lié au niveau de la séquence consensus serine-glycine dans le cas des chaînes HS/CS principalement (Esko & Zhang 1996). Une première information se situe dans la séquence primaire en acides aminés de la protéine cœur. En effet, il a été montré chez le bétaglycan, un HSPG, que la présence d'une séquence riche en acides aminés acides (aspartate, glutamate) favorisait la présence de chaîne HS (López-Casillas et al. 1994). La distance entre la séquence consensus et les acides aminés acides joue également un rôle important. Ces résultats ont également été confirmés pour le syndécan-1. La présence d'un tryptophane vicinal à la glycine de la séquence consensus ou d'autres motifs serylglycines sont également des signaux positifs dans le choix de chaîne de type HS. En l'absence de ces signaux, la formation de chaîne de type CS est préférée. Le choix dans le type de chaîne de GAG est également lié aux modifications du tétrasaccharide *linker* comme nous avons pu le voir et ces modifications pourraient être liées aussi à la séquence primaire de la protéine cœur.

b) ... dans des phénomènes de signalisation cellulaire

Certaines protéines cœurs sont également capables d'induire des signaux intracellulaires de manière plus ou moins directe. La signalisation « directe » est médiée par les protéines cœurs intégrales, c'est le cas des syndécans possédant des tyrosines conservées dans leur domaine cytosolique et pouvant déclencher la transduction de signaux au sein de la cellule. Ils sont également capables de s'oligomériser et de provoquer des événements endocytaires grâce à leurs interactions avec le réseau d'actine (Beauvais & Rapraeger 2004). Les protéines cœurs peuvent aussi interagir avec des récepteurs cellulaires : les TLR avec le biglycan, le récepteur du facteur de croissance *insulin-like* I et de l'épiderme avec la décorine (Merline et al. 2009) *etc.* Ainsi, ils initient des phénomènes de réponses immunitaires, de proliférations...

c) ... dans l'architecture matricielle

Les protéines cœurs des PG sont également des actrices incontournables de l'architecture de la MEC et de l'interaction de celle-ci avec les cellules. Les hyalectans sont prépondérants dans ces structures. En effet, ils sont capables de se lier à la fois à l'HA et aux glycoprotéines pouvant jouer le rôle de liant entre matrice et cellule. Le CD44 cellulaire relie également les cellules à leur matrice de par son interaction avec l'HA. D'autres PG modulaires, comme nous l'avons vu, sont capables d'interaction avec les intégrines cellulaires, la matrice (e.g le perlécan) *etc.*

Ainsi nous avons pu voir quelques rôles importants de la partie protéique des PG dans les phénomènes de cohésion tissulaire, mais également dans certains phénomènes cellulaires (immunité, prolifération...) et dans le choix de la chaîne de GAG qui sera liée à la protéine. Cette chaîne est également importante dans la fonctionnalisation de protéines impliquées dans une pléthore de rôles.

2. Rôles physiologiques des GAG cellulaires

Les GAG, et surtout les HS du fait de leur localisation ubiquitaire et leur très grande diversité structurale forment la pierre angulaire de nombreux processus cellulaires, notamment en fonctionnalisant et en régulant les protéines avec lesquelles ils interagissent (Figure 31). Du fait de leur nature polyanionique (groupements carboxyliques, sulfates), les GAG se lient majoritairement de manière électrostatique avec leur(s) partenaire(s), mais également grâce aux forces de van der Waals, des liaisons hydrogènes, et par des interactions hydrophobes. Ces interactions ne se font pas uniquement de manière aspécifique, certaines sont même très spécifiques (Capila & Linhardt 2002) (Kreuger et al. 2006). Des motifs d'interaction sur de nombreuses protéines ont été déterminés comme responsable de leur liaison à l'héparine (Hileman et al. 1998). Ces domaines de fixation à l'héparine (HBD) sont constitués de résidus basiques (B) (lysine, arginine et plus rarement l'histidine) pouvant être

séparés d'acides aminés quelconques (X) et agencés en motifs XBBXBX ou XBBBXXBX (Cardin & Weintraub 1989).

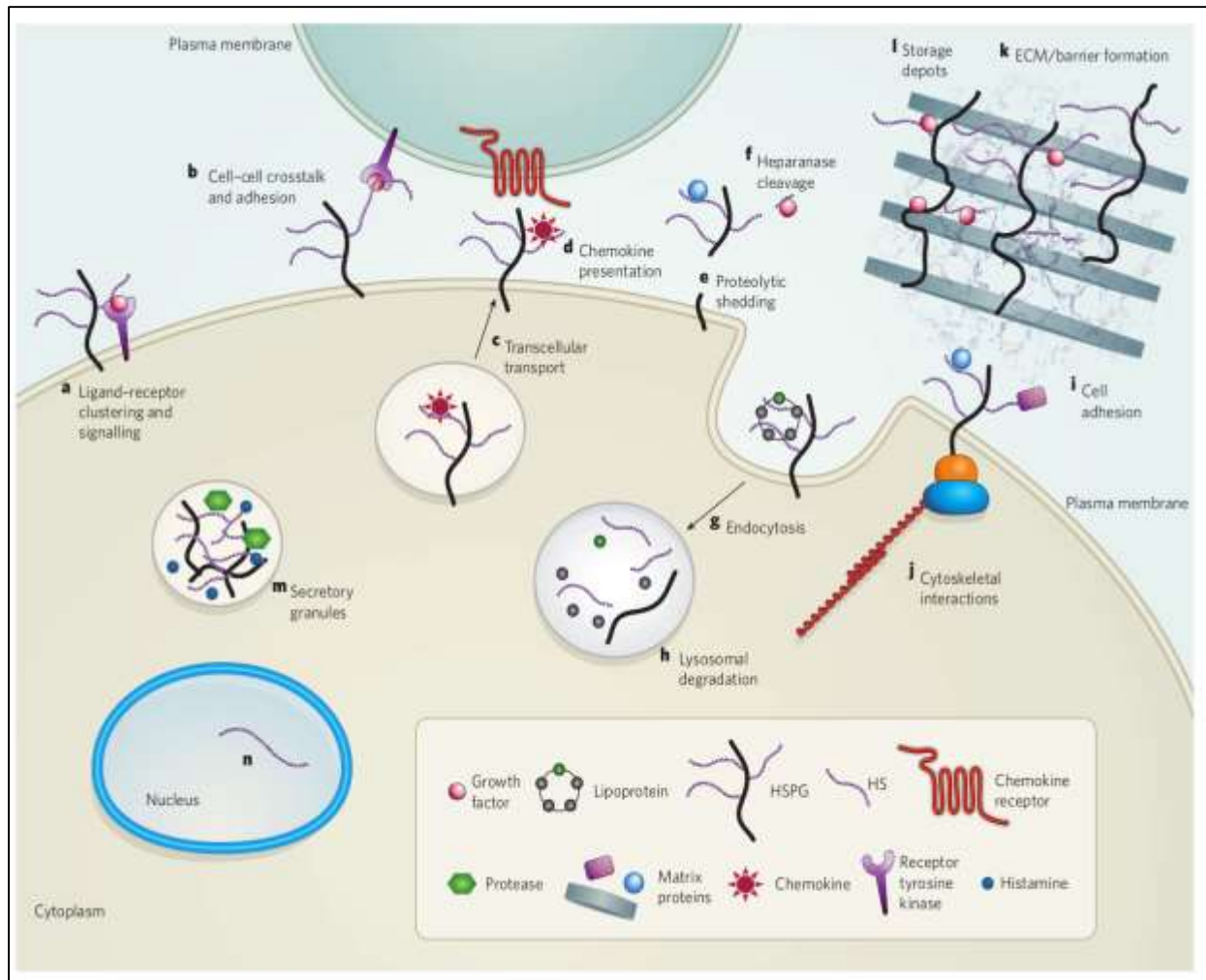


Figure 31 : Schéma des différents rôles physiologiques des protéoglycanes à héparane sulfate. Les HSPG servent de corécepteurs pour des facteurs de croissance et leurs récepteurs qui sont présents soit sur la même cellule (a) soit sur une cellule adjacente (b). Ils transportent les chimiokines à travers la cellule (c) et les présentent à la surface cellulaire (d). Les processus de shedding des syndécans et glypicans de la surface cellulaire (e) et de clivage des chaînes d'HS par l'héparanase (f) libèrent les ligands (tels que les facteurs de croissance). Les HSPG des surfaces cellulaires sont activement endocytés (g) et peuvent être recyclés à la membrane plasmique ou dégradés dans les lysosomes (h). Les HSPG facilitent aussi l'adhésion cellulaire aux MEC (i) et forment des liens avec le cytosquelette (j). Les HSPG sécrétés sont impliqués dans la formation de MEC organisées qui forment des barrières physiologiques (k) et qui séquestrent les facteurs de croissance et morphogènes pour un relargage plus tardif (l). La serglycine portant des chaînes d'héparine hautement sulfatées est contenue dans des granules de sécrétion des cellules hématopoïétiques (m). Enfin, certaines expériences suggèrent que des chaînes d'HS sont présentes dans le noyau (n) bien que leur fonction dans cet organe reste inconnue. D'après (Bishop et al. 2007).

*a) Les HS/Hp peuvent provoquer des changements conformationnels :
cas de l'héparine dans la coagulation*

Le phénomène a été bien étudié dans l'hémostase de par l'interaction d'une séquence précise d'oligosaccharide avec l'antithrombine-III (ATIII), un inhibiteur de sérine-protéase (serpine) (Jin et al. 1997). Cette protéine est présente dans le sang sous une forme non-affine pour la thrombine qui est l'enzyme finale de la cascade de réactions aboutissant à la formation

d'un caillot de fibrine. En présence d'une séquence de 5 sucres spécifiques (qui comprend une 3-O-sulfatation) retrouvée dans les HS et principalement dans les Hp (1/3 des chaînes d'Hp contiennent le motif pentasaccharidique), l'ATIII va interagir avec le motif et la protéine va subir un changement de conformation au niveau d'une boucle, exposant au solvant une arginine critique pour l'interaction avec la thrombine (Figure 32). La thrombine est également

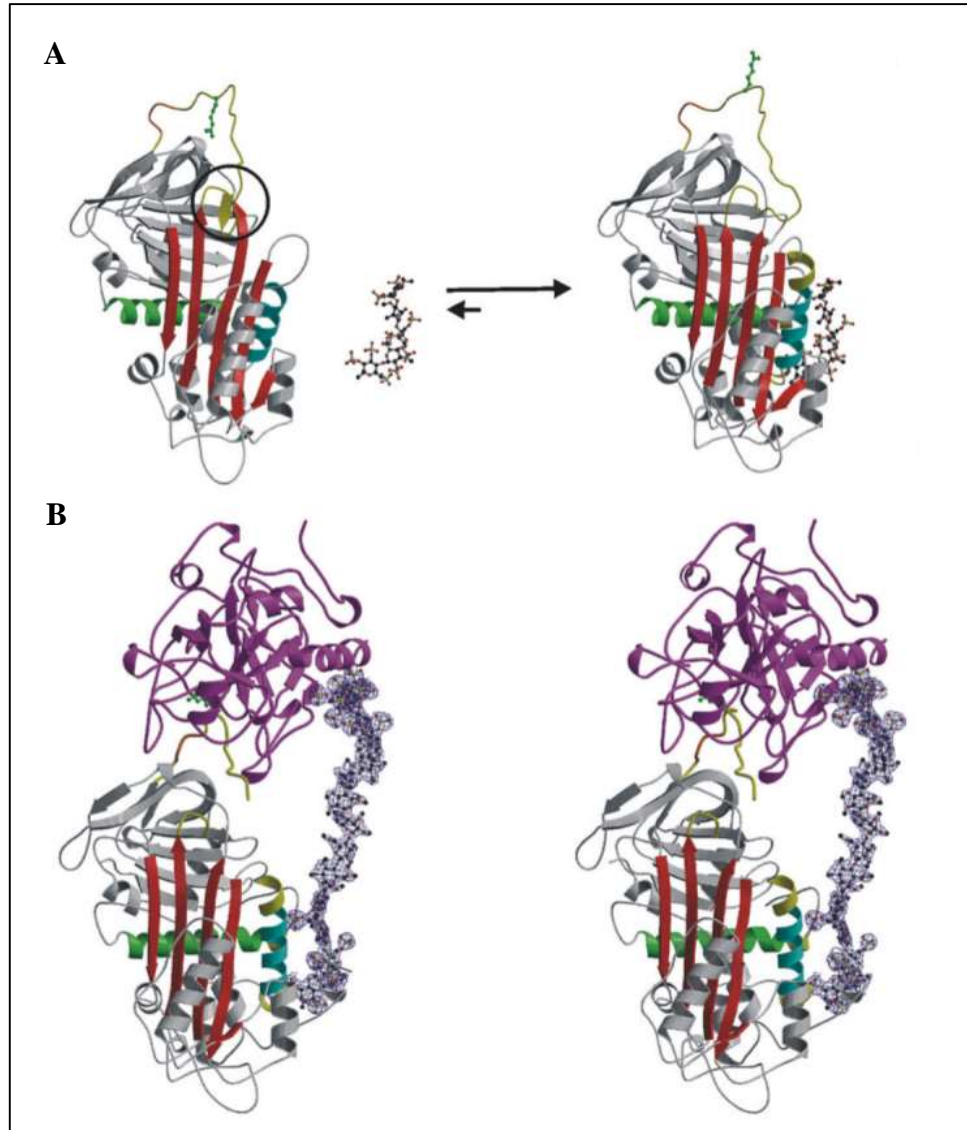


Figure 32 : Changements structuraux de l'ATIII provoqués par l'interaction avec un motif pentasaccharidique particulier de l'héparine (A). Complexe ternaire AT-III–Thrombine–Héparine (B). Structures obtenues par cristallographie aux rayons X. D'après (Li et al. 2004).

capable de se lier à la chaîne d'HS/Hp de manière moins spécifique, favorisant son interaction avec l'ATIII. La thrombine au niveau de ce trimère est totalement inhibée et le phénomène de coagulation ne peut avoir lieu (Li et al. 2004). La puissance de l'effet anticoagulant de l'héparine s'explique également par son interaction directe ou indirecte avec d'autres protéines de la voie de coagulation comme le cofacteur d'héparine II (HCII) et le facteur Xa. La formation des différents complexes aboutissent à l'inhibition de la thrombine *via* différentes voies, expliquant l'efficacité de l'héparine.

b) Les HS/CS dans la rétention et protection de protéines

Les GAG présents à la surface des cellules et au niveau de la matrice extra-cellulaire forment une zone d'interaction avec de nombreuses petites protéines comme les facteurs de croissance (Vlodavsky et al. 1991) et les cytokines (Figure 31). Ces dernières possèdent souvent des séquences d'acides aminés basiques (lysines et arginines principalement) disposées de telle manière qu'elles rendent possible l'interaction avec les GAG sulfatés comme vu précédemment. Ces médiateurs vont être ainsi concentrés au niveau de leur lieu de sécrétion par les cellules du tissu. Au niveau des mastocytes, les Hp/CS-E des serglycines contenues dans les granules concentrent également les protéases pouvant intervenir dans des phénomènes d'inflammation lors de stimuli (Humphries et al. 1999).

Une seconde fonction de l'interaction GAG-protéines est la protection de ces dernières vis-à-vis de la protéolyse. En effet, il a été montré que lorsqu'une chaîne de GAG interagissait avec le facteur de croissance FGF-2, la chimiokine CXCL12 α (ou SDF-1 α), la cytokine interféron gamma (IFN γ)... ceux-ci devenaient moins sensibles à l'action de protéases (Saksela et al. 1988) (Sadir et al. 2004) (Lortat-Jacob & Grimaud 1991). En revanche, l'action d'enzymes de dégradation des GAG comme l'héparanase permet de relarguer en solution les protéines qui leur étaient liées, et ces protéines vont alors pouvoir jouer leur rôle, notamment en se liant à des récepteurs cellulaires (Vlodavsky et al. 2002) (Figure 31). Cela va permettre aux cellules d'avoir un contrôle local et temporel sur la disponibilité en cytokines, facteurs de croissance *etc.* Ces mécanismes de sécrétion comme, dans le cas de l'héparanase, jouent également des rôles lors de phénomènes pathologiques, notamment dans la croissance anarchique de cellules tumorales.

c) Les HS comme corécepteurs

Le rôle de corécepteur des domaines NS des HS/Hp dans la signalisation de facteur de croissance a été découvert en parallèle en 1991 par deux équipes (Yayon et al. 1991) (Rapraeger et al. 1991). La famille des FGF (*Fibroblast Growth Factors*) comprend une vingtaine de membres pour 7 récepteurs cellulaires. L'un des FGF qui a été le plus étudié est le FGF-2 ou FGF basique (bFGF) jouant des rôles dans la prolifération, différenciation et migration cellulaire. Les différents FGF sont capables de se fixer sur plusieurs récepteurs au FGF (FGFR) avec des affinités différentes (Mohammadi et al. 2005) (Figure 31). Toutefois, pour déclencher une transduction de signal à travers le récepteur, celui-ci doit se dimériser, et la fixation d'une chaîne d'HS/Hp en plus du FGF est nécessaire. L'oligosaccharide impliqué dans l'interaction va permettre de rapprocher et de stabiliser les partenaires du complexe. Ce complexe est composé d'un dimère FGF:FGFR de stœchiométrie 2:2 avec une ou deux chaînes d'HS selon les modèles proposés dans la littérature (Pellegrini et al. 2000) (Schlessinger et al. 2000). Le mode d'interaction du complexe ternaire FGFR:FGF:HS/Hp fait toujours débat mais plusieurs modèles pourraient exister en fonction du type de FGF et du récepteur engagés.

La structure de la chaîne de GAG est également importante. Dans le cas de l'interaction FGF2-FGFR1, il a été montré l'importance des IdoA2S dans la fixation de la protéine sur le

récepteur et de la présence d'une glucosamine 6-*O*-sulfatée pour la fixation et l'activation efficace du FGFR1 par FGF2 (Pye et al. 1998). Cette interaction engagerait également un oligosaccharide d'une taille minimale de 10 unités osidiques. Au niveau intracellulaire, les parties cytoplasmiques des récepteurs, rapprochées lors de l'interaction, vont pouvoir s'autophosphoryler et déclencher la cascade de transduction du signal. Un modèle de cette interaction est illustré figure 33.

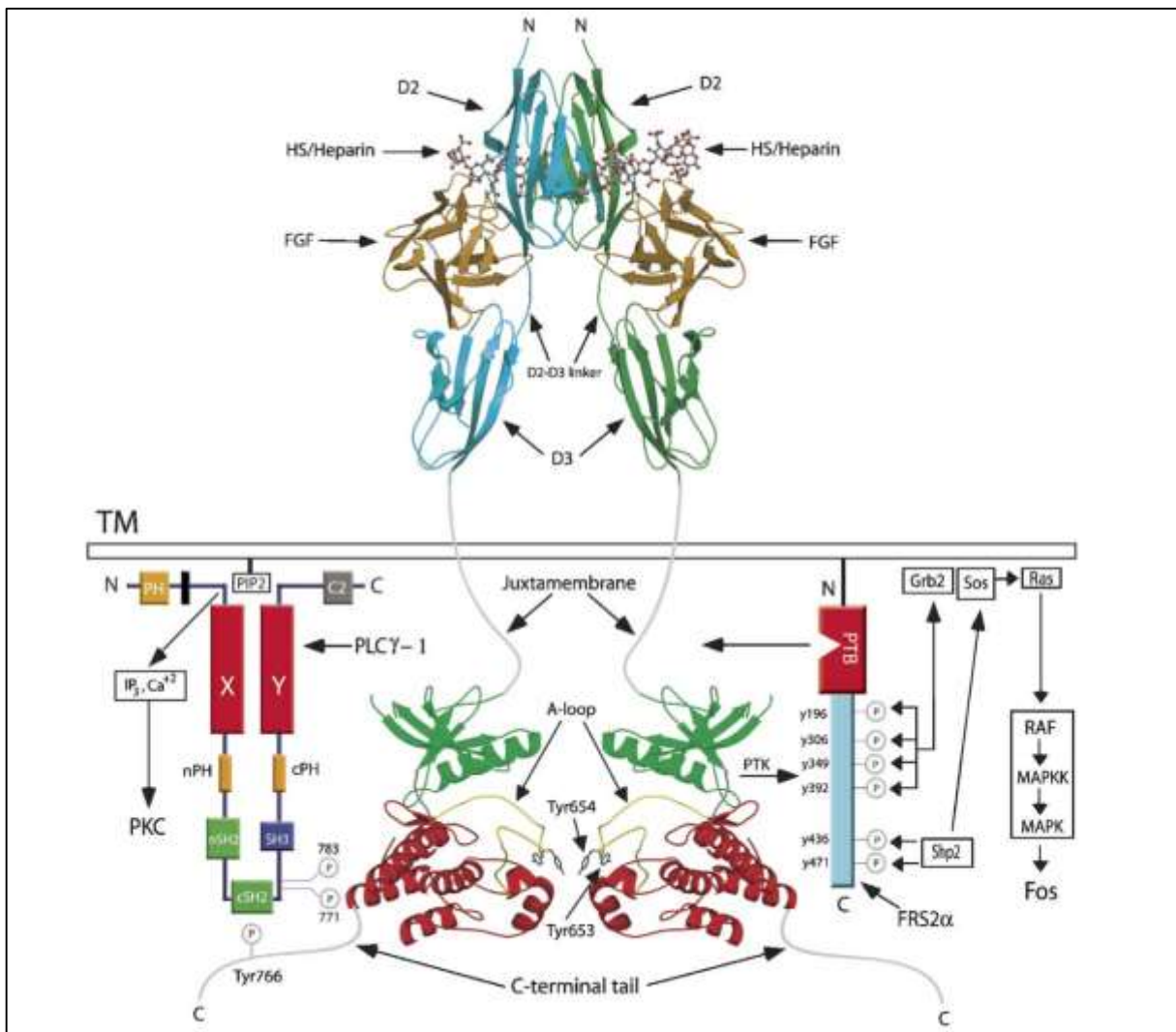


Figure 33 : Représentation schématique et structurale d'une interaction FGF-HS/Hp-FGFR et de la cascade de réactions intracellulaires. D'après (Mohammadi et al. 2005).

d) Les HS dans la formation de gradients de chimiokines et dans l'inflammation

Le recrutement leucocytaire au niveau de sites spécifiques est un mécanisme clé dans les phénomènes immunitaires en guidant les cellules nécessaires sur les sites d'inflammation, ou d'infection par exemple. Les protéines capables de recruter et guider les leucocytes sont de petites cytokines chimioattractives appelées chimiokines (*e.g* CXCL12, interleukine-8, ...). Ces petites protéines d'une dizaine de kDa possèdent des motifs basiques dans leur séquence

primaire ce qui va leur permettre d'interagir avec les HS. La fixation aux HS/Hp se ferait au niveau de clusters selon la chimiokine et sa forme oligomérique (Lortat-Jacob et al. 2002).

Lors d'un phénomène d'inflammation, les macrophages résidents du tissu vont sécréter des chimiokines et cytokines pro-inflammatoires induisant l'expression de molécules d'adhésion (type sélectines) au niveau des cellules endothéliales (Carlos et al. 1990). Les chimiokines vont également diffuser localement au niveau du site de l'inflammation et interagir avec les HS présents. Cette diffusion locale forme un gradient de chimiokines autour du site, la concentration étant de plus en plus réduite à distance du site enflammé. Par transcytose, les chimiokines vont être également présentées au niveau apical des cellules endothéliales, c'est-à-dire au niveau de la circulation sanguine, dans la lumière des vaisseaux (Middleton et al. 1997) (Middleton et al. 2002). Les leucocytes présents dans le sang vont pouvoir s'ancrer à la surface des cellules *via* les sélectines et suivre le gradient de chimiokines : c'est le phénomène de rolling. Les chimiokines interagissent également avec des récepteurs leucocytaires activant l'expression d'intégrines ce qui permet aux leucocytes de s'ancrer plus solidement à la surface de l'endothélium (Tanaka et al. 1993) (Rot 1992). S'en suit l'extravasation du leucocyte de la lumière vers la membrane basale au travers des cellules endothéliales : c'est la diapédèse. La matrice épaisse de la membrane basale va être dégradée par les leucocytes pour leur passage en sécrétant des héparanases, dégradant les HS des HSPG (perlécan, agrine...). Cette dégradation va de plus relarguer des cytokines et facteurs de croissance stockés au niveau des HS. Le leucocyte va pouvoir continuer son déplacement au niveau de cette matrice (on parle d'haptotaxie) et se diriger au niveau du site de l'inflammation où il pourra jouer ses rôles (Parish 2006) (Figure 34).

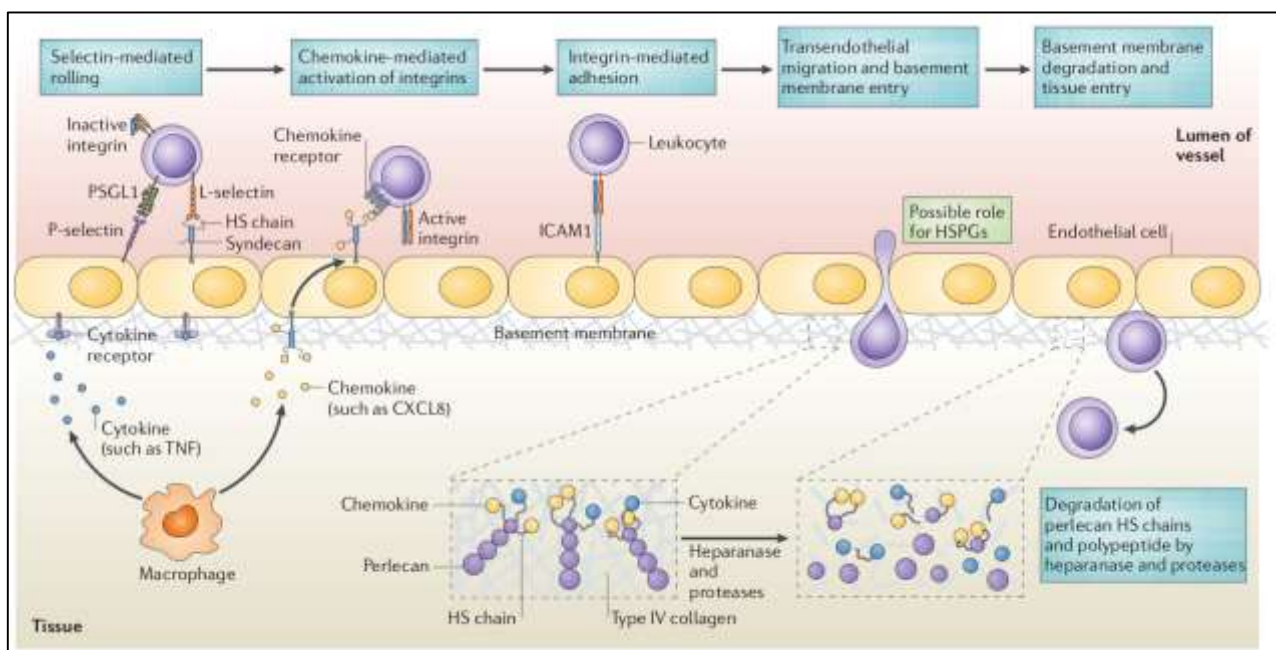


Figure 34 : Schéma d'un guidage leucocytaire vers un site inflammatoire et rôle des HSPG dans les différents stades. D'après (Parish 2006).

Ainsi, nous avons pu voir quelques rôles physiologiques essentiels des HS qui reposent sur leur interaction possible avec un grand nombre de molécules. Ces interactions permettent de stabiliser, fonctionnaliser (par changement conformationnel, par multimérisation...) les

protéines, de s'associer également avec des récepteurs et de déclencher de multiples voies de signalisation ou d'évènements cellulaires.

3. Les GAG dans des phénomènes pathologiques

La diversité structurale unique des HS/Hp et CS est également utilisée par plusieurs pathogènes afin de s'adsorber au niveau des tissus de l'organisme pour l'infecter et s'expandre.

a) GAG et infections

De nombreux virus, bactéries et parasites ont été décrits dans la littérature comme interagissant avec les chaînes de GAG présents dans les matrices de l'organisme et au niveau des surfaces cellulaires. La plupart des pathogènes n'utilisent les GAG que comme premier point d'ancrage dans l'organisme avant une interaction (en *cis* ou en *trans*) avec un récepteur membranaire qui va permettre son internalisation, et d'infecter la cellule. C'est le cas du virus de l'herpès HSV-1 (reconnaissant les glucosamines 3-O-sulfatées) *via* sa glycoprotéine D (gD), de l'hépatite C *via* la glycoprotéine 2 de l'enveloppe (E2), de l'hépatite B, du papillomavirus humain, de l'adénovirus, du virus de la dengue (Chen et al. 2008)... Le virus de l'immunodéficience humaine-1 (VIH-1) peut également interagir avec les HS *via* la glycoprotéine de 120kDa (gp120) de son enveloppe (Mondor et al. 1998). Des bactéries peuvent aussi se servir des GAG pour leur fixation au sein de l'organisme, comme *Mycobacterium tuberculosis* (Pethe et al. 2001), *Listeria monocytogenes* (Alvarez-Dominguez et al. 1997)... D'autres bactéries comme *Streptococcus pneumoniae* et *Pseudomonas aeruginosa* expriment des métalloprotéases pour induire le phénomène de *shedding* de PG. Les bactéries vont ainsi interagir avec ces fragments et se « camoufler » du système de surveillance immunitaire (Chen et al. 2007) (Park et al. 2001). Les parasites interagissent également avec les chaînes de GAG, *Toxoplasma gondii* avec les HS et CS (Carruthers et al. 2000) alors que *Plasmodium falciparum* est plus affiné pour les chaînes de CS-A (Fried et al. 2006).

b) GAG et cancer

La cancérogenèse est un phénomène complexe qui se caractérise principalement par la prolifération anarchique de cellules cancéreuses capable de vasculariser le tissu tumoral formé, d'envahir l'organisme en passant par la voie sanguine (on parle de métastases), mais également de se (dé)différencier. Ces événements (en dehors des métastases) sont normalement contrôlés dans un organisme sain et font intervenir les GAG. En effet, comme nous avons pu le voir, les GAG et notamment les HS forment un réservoir de facteurs de croissance. En destabilisant les HS, principalement par la sécrétion d'héparanase par les cellules cancéreuses (Vlodavsky et al. 2002), les facteurs de croissance contenus dans la MEC vont être libérés et vont stimuler la prolifération des cellules dont la cellule cancéreuse. De

plus, l'expression de certains HSPG et des enzymes impliquées dans la biosynthèse ou modification des chaînes d'HS est différente dans la plupart des cancers par comparaison aux cellules saines (Mythreya & Blobe 2009) (Lai et al. 2008). Cela aboutit à un nouveau profil saccharidique par rapport au tissu d'origine. Les interactions avec le milieu environnant vont donc être différentes, améliorant l'affinité pour certains facteurs nécessaires au développement tumoral. Cet environnement va également être perturbé car les cellules tumorales acidifient le pH du tissu (Gerweck & Seetharaman 1996), et cette acidité va également perturber les interactions protéines-sucres et protéines-protéines du fait des différentes protonations possibles des acides aminés.

A partir d'une certaine taille, toutes les cellules tumorales en amas n'ont plus accès aux ressources nutritives de l'environnement et vont devoir vasculariser la tumeur à partir des vaisseaux préexistant pour capter les nutriments du sang : cette néovascularisation est appelée angiogenèse. Ce processus nécessite l'interaction de facteurs de croissance avec la cellule, et notamment les VEGF (Vascular Endothelial Growth Factors) qui sont également retenus dans les HS des matrices où ils vont jouer un rôle fondamental (Harris et al. 2013). Là encore, la sécrétion d'héparanase par les cellules cancéreuses va permettre de libérer ces facteurs angiogéniques qui vont stimuler les récepteurs cellulaires. L'héparanase joue également un rôle dans la surexpression de VEGF par la cellule, stimulant davantage la vascularisation du tissu tumoral (Zetser et al. 2006).

Un des derniers phénomènes dans la maladie est le détachement des cellules et leur colonisation de l'organisme à travers le courant circulatoire : c'est le phénomène de métastase. Pour cela, les cellules doivent remodeler leur surface cellulaire ou sécréter des enzymes de dégradation pour diminuer ou abolir les interactions entre cellules ou avec la matrice. Les HA, CSPG et HSPG constituant les MEC et les surfaces cellulaires jouent des rôles essentiels dans l'adhésion des cellules et leur cohésion au sein du tissu. Le changement de profil en GAG (principalement des HS) peut affaiblir ces interactions intercellulaires et cellules-matrice facilitant ainsi la migration des cellules cancéreuses. Le *shedding* ou le clivage des GAG des PG cellulaires abolit de manière efficace ces liaisons et permet l'invasion des cellules cancéreuses dans les tissus voisins. L'intravasation et l'extravasation de celles-ci peuvent être reliées aux HS car les cellules cancéreuses sont capables de chimiotaxie, et les HS de former des gradients de chimiokines comme nous l'avons vu pour les phénomènes immunitaires (van Zijl et al. 2011). L'efficacité de cette chimiotaxie est accrue par la surexpression des récepteurs aux chimiokines (Müller et al. 2001).

Enfin, les cellules cancéreuses sont généralement des cellules qui sont revenues à un stade de cellule non-différenciée. Ces cellules expriment à leur surface différentes protéines, dont les HSPG, et dans des proportions différentes par rapport à la cellule différenciée dont elle est issue (Knelson et al. 2014a).

Tous les HSPG n'interviennent pas de la même manière dans les différents stades de cancers et dans les différents cancers, deux HSPG pouvant même avoir des effets antagonistes sur le développement tumoral. Ces effets viendraient de la structure de la chaîne d'HS et de sa

capacité à interagir avec des facteurs différents. Les rôles majeurs des HSPG sont schématisés figure 35.

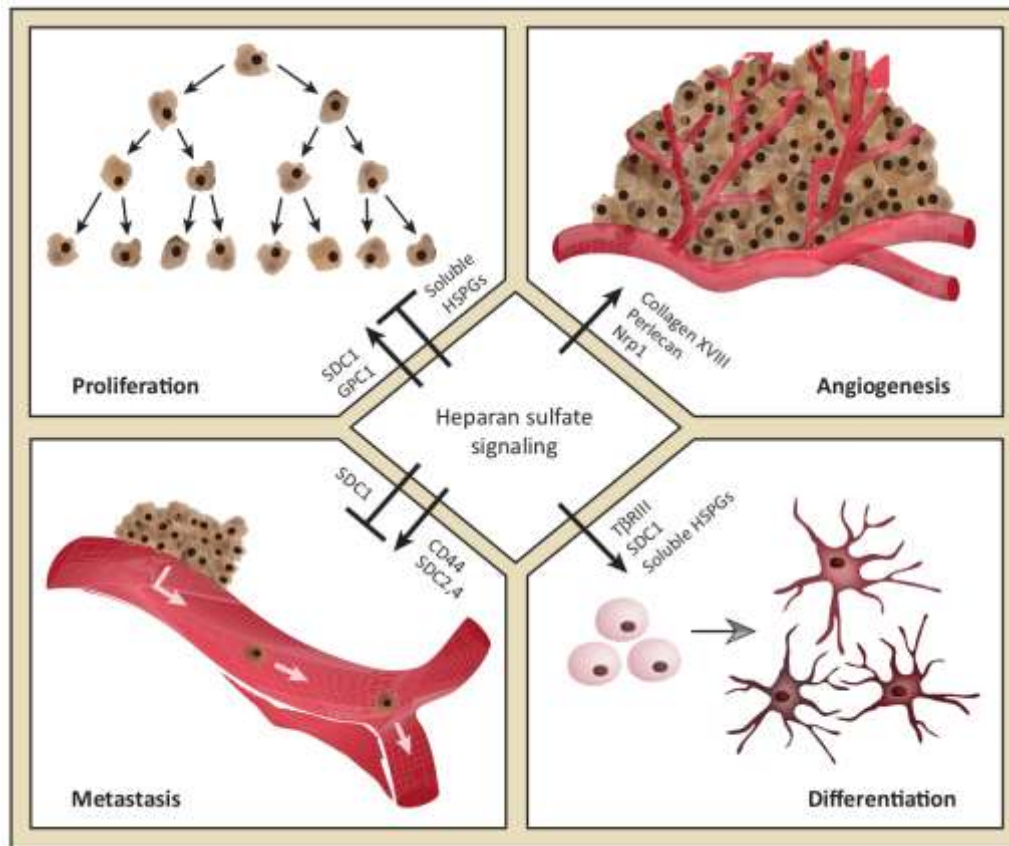


Figure 35 : Effets des protéoglycanes à héparane sulfate dans les différentes étapes du cancer. SDC : syndecan ; GPC : glypican ; TGFβIII : type III transforming growth factor β receptor ; Nrp1 : neuropiline. D'après (Knelson et al. 2014b).

c) Les maladies génétiques

Les maladies génétiques liées à un dysfonctionnement de la biosynthèse, modification ou catabolisme des GAG sont extrêmement nombreuses et variées. Comme nous avons pu le voir jusqu'à présent, des mutations génétiques peuvent aboutir à la non-polymérisation de la chaîne de GAG ou à des défauts de structures pouvant être létales ou entraîner des malformations variées. Les défauts d'enzymes de dégradation des GAG sont également des pathologies héréditaires rares appelées mucopolysaccharidoses (MPS) et engagent le pronostic vital de la personne atteinte. Pour revue voir Sarrazin *et al* 2011 et figure 30.

E. Intérêts de l'étude des GAG

Comme nous avons pu le voir précédemment, les PG sont impliqués dans de nombreux phénomènes physiologiques, et peuvent être dérégulés lors de comportements pathologiques. La grande variété de structure des GAG (HS/Hp et CS), leur propriété rhéologique (HA...) ont permis d'envisager les GAG comme médicaments potentiels et marqueurs de maladie. Historiquement, l'intérêt des GAG en tant que médicament a

commencé avec la découverte du pouvoir anticoagulant de l'héparine au début du XXe siècle par McLean (1916) qui a rapidement été utilisé comme antithrombotique chez l'Homme. Dans cette partie, nous verrons quelques applications des GAG en tant que médicament et leur importance dans l'industrie d'aujourd'hui et leur potentiel pour le futur.

1. Les GAG comme médicaments

a) Dans l'hémostase

Comme décrit précédemment, l'héparine en se liant à l'ATIII, induit un changement conformationnel de la protéine, la rendant ainsi capable d'interagir avec la thrombine, enzyme responsable de la formation de fibrine, et de l'inhiber. L'interaction se fait *via* un motif pentasaccharidique précis principalement retrouvé dans les chaînes d'Hp et dans une moindre mesure dans les chaînes d'HS. De même, le dermatane sulfate (DS) a montré également des propriétés anticoagulantes. Son activité n'est pas liée à l'ATIII car il ne peut pas interagir avec elle du fait de sa structure saccharidique différente : présence de galactosamines, absence de *N*- et 3-*O*-sulfatation... L'activité anticoagulante des DS provient de son interaction avec le cofacteur d'héparine II (HCII) qui est un homologue de l'ATIII. Ce complexe binaire va permettre d'augmenter la vitesse d'interaction avec la thrombine d'un facteur 1000, inhibant l'activité de cette dernière (Tollefsen 2010).

b) Contre l'arthrose

L'arthrose est la forme d'arthrite la plus répandue. Elle résulte d'une dégradation du cartilage situé entre deux os qui peuvent alors être en contact direct causant des douleurs articulaires. L'origine de cette érosion du cartilage est principalement mécanique, c'est pour cela que les personnes souffrant d'arthrose sont généralement des personnes de plus de 60 ans ou en surpoids.

L'acide hyaluronique est un des composants essentiels des cartilages articulaires. Grâce à son fort pouvoir rhéologique, il permet de lubrifier ceux-ci et de protéger les os des stress mécaniques. Une réduction en HA est observée dans l'arthrose. Des injections en HA commercial ont montré des effets bénéfiques temporaires chez les patients atteints d'arthrose au niveau des genoux (Wobig et al. 1999). L'HA est toujours utilisé aujourd'hui pour épaissir le cartilage au niveau des articulations mais cela ne représente qu'une solution temporaire, nécessitant des injections régulières toute la vie. L'efficacité de ce type de traitement pour l'arthrose a été récemment questionné en raison de conflits d'intérêts que pourraient avoir certains auteurs des études publiées dans la littérature (Printz et al. 2013).

c) Comme anticancéreux

Les cibles dans la lutte contre les cellules cancéreuses sont multiples. En effet, le cancer peut se diviser en plusieurs phases comme la croissance tumorale, l'angiogenèse et la

formation de métastases. Dans chacun de ces stades du cancer, l'héparine a montré son efficacité et est aujourd'hui utilisé dans certains traitements (Fuster & Esko 2005).

Lors de la croissance tumorale, des enzymes de dégradation comme l'héparanase sont sécrétées dans les MEC. Cela permet le relargage dans le milieu de facteurs de croissance retenus par les HS et donc de stimuler la croissance cellulaire. Bien que structuralement proche, l'héparine n'est pas un bon substrat des héparanases contrairement aux HS. L'héparine va cependant se fixer à l'enzyme, mais sans être digérée, bloquant ainsi le site actif de l'enzyme (Bar-Ner et al. 1987) (Nakajima et al. 1988). Cette inhibition de l'héparanase par l'héparine va empêcher, tout du moins en partie, la dégradation des HS (par compétition) et donc du relargage des facteurs protéiques interagissant avec ces derniers. Parmi ces facteurs se trouvent aussi des facteurs angiogéniques (comme les VEGF), faisant ainsi de l'héparine également un inhibiteur de l'angiogenèse.

Au niveau de stades plus tardifs, comme lors de métastases, l'héparine, selon sa source, a montré une action bénéfique en allongeant de plusieurs mois l'espérance de vie des patients (Tagalakakis et al. 2007). Lors de métastases, les cellules cancéreuses passent dans le flux sanguin et peuvent interagir avec des cellules endothéliales en formant des tumeurs loin du site primaire. Ces cellules cancéreuses peuvent également interagir avec d'autres cellules dans le sang, comme les leucocytes, ou des plaquettes et former des amas pouvant conduire à des embolies. L'interaction entre cellules cancéreuses et composants sanguins est due aux interactions entre des glycoprotéines (comme les mucines...) de la cellule cancéreuse avec les sélectines des cellules endothéliales, plaquettes et leucocytes. L'héparine étant capable d'interagir avec les sélectines, va entrer en compétition avec les mucines. Les cellules cancéreuses ne vont plus pouvoir adhérer, ou dans une moindre mesure selon la quantité d'héparine dans le sang, et vont mourir (Stevenson et al. 2005). Ainsi l'héparine joue un rôle bénéfique dans l'arrêt de la progression métastatique des cellules cancéreuses. Toutefois, du fait de ses propriétés anticoagulantes, l'héparine peut induire des effets secondaires comme des hémorragies. Son utilisation et son dosage sont donc au cas par cas. De nouvelles molécules sont à l'étude avec un squelette d'héparine mais dépourvues de certaines sulfatations pour abolir l'interaction avec l'ATIII tout en préservant leur capacité d'interagir avec les sélectines et les héparanases (Borsig et al. 2011).

En résumé, les GAG et dérivés sont des molécules de choix dans la lutte contre diverses pathologies et infections grâce à leurs propriétés d'interaction et rôles très divers. Les GAG, et plus généralement les PG peuvent également être des cibles dans des maladies où ceux-ci sont au centre de dysfonctionnements comme lors de leur accumulation dans les mucopolysaccharidoses. *In vitro* l'ajout des enzymes de dégradation permet de retrouver un phénotype sain (Daniele et al. 2002). Ces résultats ont ouvert la voie à des thérapies de remplacement enzymatique (ERT) déjà à l'étude *in vivo* (Higuchi et al. 2012) et dont certaines sont déjà disponibles sur le marché (Wraith et al. 2008). L'inhibition d'interaction des GAG, et principalement des CS, en neurobiologie est également un terrain d'investigation intéressant pour la guérison de paralysies et de traumatismes en restaurant la régénération axonale normalement inhibée par les CSPG chez l'Homme (Moon et al. 2001) (Bradbury et al. 2002).

Les GAG ou leurs dérivés, seuls ou combinés avec des peptides, sont des molécules d'intérêt croissant pour combattre les phénomènes d'infections et notamment d'entrée virale, laissant imaginer tout leur potentiel en tant que médicament (préventif ou curatif). Un exemple récent est la synthèse d'une molécule hybride au sein du groupe « Structure et Activité des Glycosaminoglycanes » de l'Institut de Biologie Structurale (Grenoble) montrant un fort potentiel d'inhibition du HIV-1 aux récepteurs de surface comparé aux médicaments déjà présents sur le marché (Connell et al. 2012).

2. Industrie des GAG et PG dans la Santé

a) L'industrie de l'héparine

Sur le marché du médicament, plusieurs préparations à base d'héparine appartenant à différents groupes pharmaceutiques sont en concurrence. On peut séparer l'héparine non-fractionnée (UFH) des héparines à faible poids moléculaire (LMWH) où on retrouve principalement le dalteparine (Fragmine® par Pfizer Inc.), le tinzaparine (Innohep® par Leo pharmaceutical company) et l'énoxaparine (Lovenox® par Sanofi). Ce dernier est le plus vendu des LMWH et même de tous les anticoagulants disponibles sur le marché en 2008 (Melnikova 2009), comprenant d'autres molécules comme les warfarines (antagonistes de la vitamine K impliquée dans l'activation de protéines clés dans la coagulation), et d'autres antagonistes des protéines hémostatiques. Les LMWH représentaient un marché de 4,5 milliards de dollars (américains) en 2008 mais tend à décliner avec le développement d'antagonistes dans la voie de coagulation présentant moins d'effets secondaires (comme des thrombocytopénies) et/ou étant plus facile à administrer (e.g par voie orale) (Figure 36). Selon l'*Intercontinental Marketing Services Health* (IMS Health), institut américain reconnu pour leurs études des médicaments, l'énoxaparine au dernier trimestre 2013 est le 37^e médicament le plus vendu, représentant environ 1,4 milliard de dollars de vente annuelle pour quasiment 17 millions d'unités.

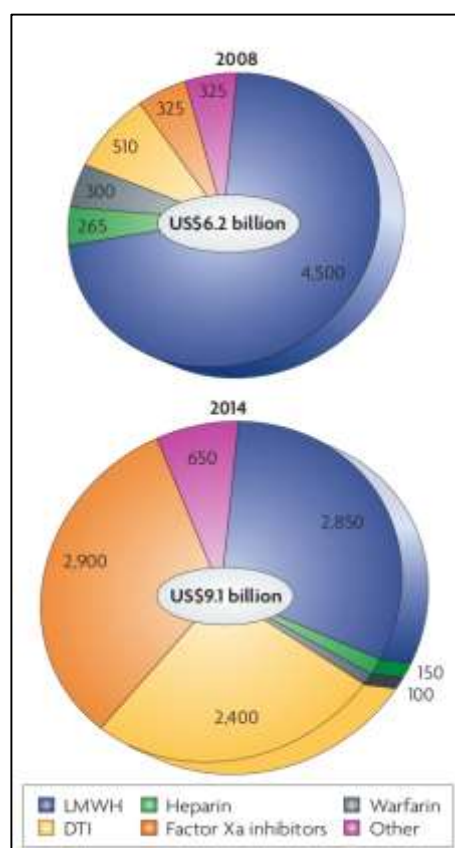


Figure 36 : Marché mondial de l'anticoagulant en 2008 (haut) et prévisions pour 2014. DTI : inhibiteur direct de la thrombine. D'après (Melnikova 2009).

Aujourd'hui, la source mondiale d'héparine comme médicament est exclusivement issue d'intestin de porc des chaînes agro-alimentaires. L'héparine de source bovine est utilisée dans la Recherche académique mais ne fait plus partie des préparations médicamenteuses à cause de contaminants potentiels pathogènes pour l'Homme. Ainsi, depuis la crise de « la

vache folle » en 1996, les héparines d'origine bovine sont interdites en Europe et en Amérique en raison de la présence potentielle de prions responsables d'encéphalopathies spongiformes bovines. La Chine est le principal producteur d'héparine avec quasiment la moitié de la production mondiale qui est autour des 100 tonnes (Linhardt & Liu 2012). En 2008, des lots frelatés d'héparine ont été découverts causant la mort de plus de 80 personnes et de plus de 700 rapports d'accidents : c'est le scandale de l'OSCS (pour *oversulfated chondroitin sulfate*).

b) Le scandale de l'OSCS (oversulfated chondroitin sulfate)

Entre 2007 et 2008, la prise d'héparine de certains lots a engendré des effets secondaires tels qu'une baisse de la tension artérielle, des nausées et vomissements, des troubles respiratoires... a reporté la *Food and Drug Administration* (FDA). Les lots ont été retirés du marché et analysés parallèlement en laboratoire. C'est en mars 2008 que la structure du contaminant est découverte : il s'agit de chondroïtine sulfate « sursulfatée » (OSCS). Les investigations remontent à la source de la contamination basée en Chine dans la société pharmaceutique étatsunienne Scientific Protein Laboratories. Cette société fournit l'héparine non-fractionnée (UFH) à d'autres compagnies qui préparent leur LMWH à partir de celle-ci. Cependant, au cours du processus de fragmentation et dépolymérisation de l'UFH, l'OSCS ne réagit pas de la même manière aux traitements, voire y est insensible (Figure 37). Ainsi malgré les différentes étapes de production, l'OSCS est toujours présent en fin de procédé, et selon les lots, a pu représenter jusqu'à 60% de la masse du médicament.

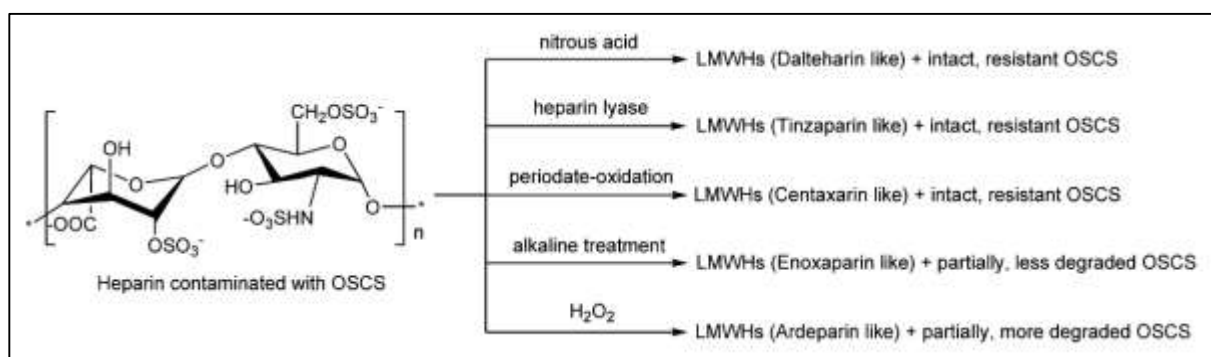


Figure 37 : Stabilités des contaminants au travers des différents procédés de préparation des LMWH. D'après (Liu et al. 2009).

La question de l'origine de cet OSCS a alors été posée, d'autant plus que les mammifères ne synthétisent pas d'OSCS. Ce GAG est retrouvé chez certains mollusques et organismes marins et son extraction, ou sa synthèse à partir de CS est bien moins onéreuse que l'héparine. De plus, il existe une production d'OSCS qui est utilisée comme nutraceutique contre l'arthrose. Il est donc apparu clairement pour la FDA que cette contrefaçon d'héparine est intentionnelle et avait pour but de vendre autant (en masse) « d'UFH » (coupée à l'OSCS) pour un coût de production final nettement moindre (Food and Drug Administration).

Ces événements ont mis en lumière la nécessité de développer des méthodes d'analyse fines et systématiques des lots d'héparine destinés à la santé humaine. En effet, OSCS et Hp sont des molécules proches de par leur poids moléculaire, leur charge nette, et leur structure. Parmi les techniques utilisées, la résonance magnétique nucléaire (RMN) peut être une méthode d'analyse de choix car elle est rapide et fournit des informations à la fois qualitative et quantitative, permettant ainsi de distinguer des lots contaminés de lots purs (Figure 38) (Liu et al. 2009).

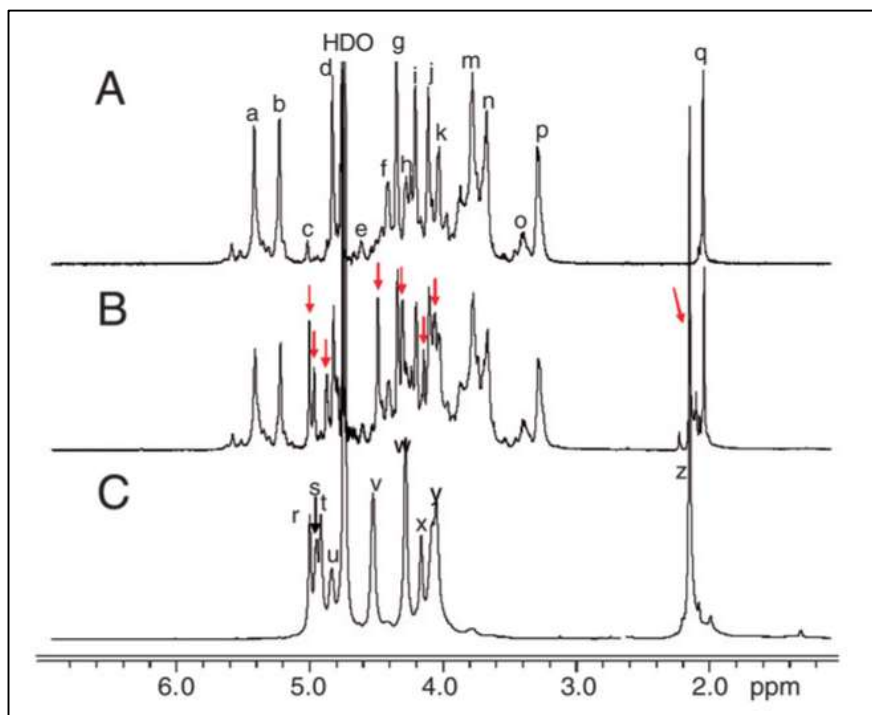


Figure 38 : Spectre RMN ^1H et attribution de spectres d'héparine pur (A), contaminé à l'OSCS (B) et d'OSCS pur (C). a : H_1 GlcNS, GlcNS6S ; b : H_1 IdoA2S ; c : H_1 IdoA ; d : H_5 IdoA2S ; e : H_1 GlcA ; f : H_6 GlcNS6S ; g : H_2 IdoA2S ; h : H_6 GlcNS6S ; i : H_3 IdoA2S ; j : H_4 IdoA2S ; k : H_5 GlcNS6S ; l : H_6 GlcNS ; m : H_4 GlcNS6S ; n : H_3 GlcNS, GlcNS6S ; o : H_2 GlcA ; p : H_2 GlcNS6S ; q : CH_3 des acétyles ; r : H_4 GalNAc2S4S6S ; s : H_3 GlcA2S3S ; t : H_1 GlcA2S3S ; u : H_1 GalNAc2S4S6S ; v : $\text{H}_{2,4}$ GlcA2S3S ; w : H_6 GalNAc2S4S6S ; x : H_5 GlcA2S3S ; y : $\text{H}_{2,3,5}$ GalNAc2S4S6S ; z : CH_3 des acétyles des GalNAc2S4S6S. Les pics avec des flèches rouges montrent des pics non-observés dans l'héparine.

c) L'industrie de l'acide hyaluronique

L'industrie de l'acide hyaluronique, le GAG le plus simple de la famille, représente un secteur à haute valeur ajoutée car il trouve à la fois des applications dans le domaine pharmaceutique et en cosmétologie grâce à ses propriétés rhéologiques. Dans le domaine du Bien-Être, l'HA est utilisé dans les crèmes à visée antiride car il est censé restaurer/améliorer, en pénétrant dans la peau, les propriétés viscoélastiques du tissu perdues avec l'âge. La société Contipro® est l'un des fabricants leader d'HA dans le domaine de la Santé et du Bien-Être en dominant 30% du marché mondial de l'acide hyaluronique. Le marché de ce viscosupplément représentera 2,5 milliards de dollars à l'horizon 2017 selon le rapport du Groupe de Recherche Millenium (MRG). L'HA, bien que naturellement présent chez les

vertébrés, est principalement extrait de bactéries *Streptococcus spp.* (Yamada & Kawasaki 2005). Ces organismes synthétisent ce polymère pour s'encapsuler à l'intérieur afin d'éviter d'être reconnu par le système immunitaire de l'hôte (Wessels et al. 1991). L'HA bactérien est plus facile à produire en fermenteur et sa purification moins coûteuse qu'une extraction à partir de tissus animaux.

Les HA trouvent également d'autres applications dans le domaine de la médecine où ils représentent des molécules d'intérêt dans le développement de nouveaux traitements. Divers types d'HA de poids moléculaires différents sont utilisés :

- Les HA de haut poids moléculaires (HMWHA pour *High Molecular Weight Hyaluronic Acid*) (supérieurs à 500kDa) : utilisés dans les chirurgies et particulièrement celle de l'œil car ils facilitent le processus de cicatrisation. Ils sont utilisés aussi comme support dans les régénérations hépatiques, et comme nous l'avons vu précédemment, pour lutter contre l'arthrose (en injections intra-articulaires).

- Les HA de faible poids moléculaire (LMWHA pour *Low Molecular Weight Hyaluronic Acid*) (10-500kDa), et oligosaccharides : induisant des phénomènes cellulaires différents, voire antagonistes au HA de haut poids moléculaire (Fuchs et al. 2013), peuvent trouver des applications dans les traitements anticancéreux. En effet, les LMWHA vont entrer en compétition avec les HMWHA pour se lier au CD44, la substitution des HMWHA par les LMWHA va rendre le réseau matriciel au niveau de la tumeur moins dense et les cellules malades vont donc être plus accessibles aux molécules thérapeutiques possiblement couplées aux HA (Luo & Prestwich 1999) (Misra et al. 2011). Toutefois, un effet secondaire indésirable lié à ce relâchement du réseau de la MEC est la favorisation de métastases en facilitant les migrations cellulaires (Kobayashi et al. 2002).

Les GAG forment donc une famille de molécules représentant un intérêt déjà démontré et toujours croissant pour subvenir aux besoins de l'Homme dans les domaines de la Santé et du Bien-Être. Nous avons pu voir jusque-là quelques implications et applications biologiques dans divers phénomènes physiologiques et pathologiques. Celles-ci reposent majoritairement sur leurs interactions avec la myriade de partenaires possibles. Une connaissance structurale de ces GAG est donc essentielle pour comprendre ces interactions *in vivo* et s'en servir pour concevoir les médicaments de demain. Dans la partie suivante, nous passerons en revue les principales méthodes d'analyse structurale et dynamique utilisées jusqu'à maintenant.

F. Etudes structurales des GAG

Les analyses structurales des GAG ont pour but la détermination de l'enchaînement des unités monosaccharidiques (longueur, modifications), leur(s) conformation(s) tridimensionnelle(s), et leur dynamique, seuls ou lors d'interaction avec un partenaire. Ces études sont rendues compliquées par la nature répétitive des GAG et nécessite très généralement leur dépolymérisation en unités disaccharidiques ou oligosaccharidiques. De plus, les séquences des GAG ne sont pas codées génétiquement comme le sont les protéines,

ce qui rend leur amplification pour analyse impossible. En effet, la biosynthèse des GAG apparaît *a priori* comme assez aléatoire et dépendante de l'expression et de l'efficacité des enzymes qui les synthétisent. Leur expression est pourtant finement régulée et elles possèdent des motifs de reconnaissance précis pour leurs partenaires. Ces motifs, parfois rares, ont pourtant une activité biologique essentielle, ce qui rend la caractérisation des chaînes primordiale. Au final, les cellules expriment et sécrètent des GAG fortement hétérogènes et potentiellement riches en information, plus que l'ADN/ARN et les protéines. C'est d'autant plus vrai pour les HS dont le motif disaccharidique peut être modifié de 48 manières différentes.

1. Détermination de la structure primaire des GAG

a) La dépolymérisation des HS

Comme nous l'avons vu précédemment, deux méthodes de dépolymérisations sont principalement utilisées en laboratoire :

(1) La dépolymérisation chimique

Principalement effectué en présence d'acide nitreux, la dépolymérisation chimique permet de cliver la liaison osidique des GlcNS côté réducteur des HS. Cette dégradation se fait à pH acide et est indépendante des autres modifications de la chaîne, seule la présence de groupements *N*-acétyles empêche la réaction. Suite à cette dépolymérisation, la glucosamine est désaminée et forme un anhydromannose (Shively & Conrad 1976). Pour pouvoir dépolymériser les domaines *NA* en disaccharides, une hydrazinolyse doit être réalisée au préalable car elle permet d'enlever les groupements acétyles. Les différents produits issus de la dépolymérisation précédente sont séparés et caractérisés par chromatographie que nous verrons par la suite (Bienkowski & Conrad 1985).

(2) La dépolymérisation enzymatique

Suite à la découverte et au clonage d'enzymes bactériennes de dégradation des HS/Hp, leur utilisation est vite devenue un outil essentiel pour l'analyse des HS. Trois enzymes majeures ont été découvertes chez *F. heparinum* (maintenant *Pedobacter heparinus*) nommées héparinases (ou héparitinases) I, II et III (Hep I-III). Les héparinases sont des lyases/éliminases qui clivent les liaisons osidiques entre les glucosamines et les acides uroniques en laissant une 4,5 insaturation sur l'acide uronique. Cette insaturation est fondamentale pour la purification des oligosaccharides par chromatographie car elle absorbe spécifiquement à une longueur d'onde de 232nm. Par contre, la formation de la double liaison en 4-5 se fait aux dépens de l'information d'épimérisation (Linker & Sampson 1960).

Chaque héparinase (Hep) a un spectre de dégradation relativement spécifique. L'héparinase I (ou héparitinase III) dégrade les liaisons osidiques entre les GlcNS et les IdoA2S, donc dans les domaines NS. A l'inverse, l'héparinase III (ou héparitinase I) coupe les liaisons entre glucosamines (N-Ac ou NS) et acides glucuroniques, donc au niveau des régions N-acétylées et intermédiaires NS/NA. L'héparinase II quant à elle a un spectre de digestion plus large, possédant 2 sites catalytiques distincts ressemblant aux sites Hep I et Hep III, lui permettant de couper les chaînes d'HS/Hp de façon aspécifique (Figure 39). Ainsi, des oligosaccharides appartenant spécifiquement aux domaines NS ou NA vont pouvoir être générés et purifiés, et servir par la suite à des analyses d'interactions et/ou caractérisation. Une digestion complète de la chaîne d'HS/Hp en combinant les trois héparinases va permettre d'aboutir à une somme de disaccharides qui constituaient les chaînes. Leur analyse ensuite par chromatographie par exemple permet de retrouver leur profil de modification et longueur (la technique est appelée analyse disaccharidique).

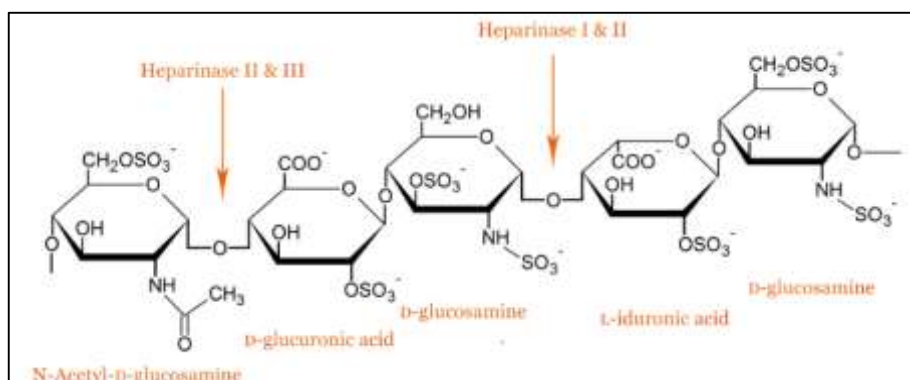


Figure 39 : Différentes spécificités de reconnaissance et de clivage au niveau des HS par les héparinases I, II et III de *P.heparinus*.

Des enzymes bactériennes spécifiques des CS/DS existent également et sont issues de *P. heparinus* pour les chondroïtinases AC I et B (dégradant les CS-A, CS-C et CS-B ou DS). *Arthrobacter aurescens* exprime les chondroïtinases AC II, quant à *Proteus vulgaris*, cette bactérie produit une chondroïtinase à spectre de digestion très large dite ABC qui dégrade les CS-A, CS-C et DS. *Pseudomonas sp* et *Bacillus sp* sont également capables de produire des enzymes de dégradation spécifiques des KS I et II. Les hyaluronidases utilisées *in vitro* proviennent principalement de *Staphylococcus aureus* et *Streptococcus pyogenes*.

b) Etudes par électrophorèse

(1) Etudes par électrophorèse capillaire (CE)

La CE est une technique d'électrophorèse qui va permettre de séparer les espèces chimiques selon leur rapport charge/taille au sein d'un petit tube appelé capillaire contenant une solution d'électrolyte et placé dans un courant électrique (Figure 40). Cette technique est utilisée pour la séparation d'oligosaccharides et disaccharides issus de GAG (Desai et al. 1993) et permet de rester en solution. De très faibles quantités d'échantillon de l'ordre du femtomole (10^{-15} mole ou fmole) voire de l'attomole (10^{-18} mole) dans les techniques CE hautes performances (Karamanos et al. 1996), peuvent être détectées. C'est une technique rapide

(séparation en moins d'une heure) et résolutive qui permet de quantifier et distinguer les différentes espèces disaccharidiques, même les plus mineures au sein d'une solution hétérogène. Toutefois, plus la chaîne de GAG est longue, plus elle va être difficile à séparer car elle va co-migrer avec d'autres oligosaccharides. La résolution de la technique est limitée et est généralement utilisée pour des tailles inférieures ou égales à l'octasaccharide (Liu et al. 2002).

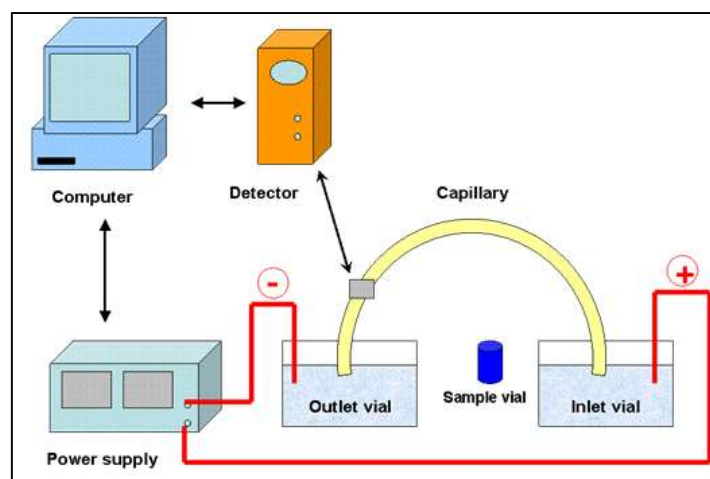


Figure 40 : Schéma d'un appareillage d'électrophorèse capillaire.

(2) Etudes par FACE (*Fluorescence-Assisted Carbohydrate Electrophoresis*)

Cette technique repose sur la digestion extensive des polysaccharides en disaccharides et à l'utilisation d'un fluorophore (le 2-aminoacridone ou AMAC) qui se lie covalamment à l'extrémité réductrice de ceux-ci. Les différentes espèces disaccharidiques sont séparées par électrophorèse sur gel de polyacrylamide. Les différents disaccharides sont révélés par leur fluorescence et la quantité de celle-ci est reliée à la quantité en disaccharide. Cette technique présente l'avantage d'être indépendante de la sulfatation des différentes espèces disaccharidiques (applicable pour tout type de GAG donc) et peut détecter une quantité de l'ordre du picomole (Calabro et al. 2001).

c) Etudes par chromatographie

Les analyses chromatographiques des GAG, et principalement des HS reposent majoritairement sur une séparation fine des espèces par chromatographie liquide haute-performance (HPLC) à travers une colonne échangeuse d'anion (SAX) (Figure 41). Les différents oligosaccharides formés par digestion enzymatique (disaccharides, tétrasaccharides, ...) sont élués en fonction de leur force d'interaction électrostatique avec la phase par différentes concentrations salines et sont détectés en sortie grâce à l'absorption de l'insaturation du $\Delta 4,5$ -uronate à 232nm. D'autres techniques peuvent être utilisées comme le marquage radioactif des oligosaccharides au niveau de leur extrémité réductrice par réduction au tétrahydroborate de sodium (ou borohydrure de sodium) tritié ($[^3\text{H}]\text{-NaBH}_4$) et comptage

de la radioactivité en sortie de colonne ; et plus récemment par couplage des oligosaccharides avec un chromophore détecté ensuite par spectrométrie UV-visible (Skidmore et al. 2010).

Le couplage de digestions enzymatiques et/ou chimiques aux séparations chromatographiques successives permettent d'aboutir à la séquence de l'oligosaccharide de départ (Pye et al. 1998) (Vivès et al. 1999). Cela permet également de créer des banques d'oligosaccharides de motifs et longueurs variés. L'utilisation de ces banques dans l'interaction avec un partenaire protéique permet de mettre en évidence les séquences oligosaccharidiques les plus affines et les motifs critiques pour les interactions. Par exemple, l'utilisation d'une banque de dp6-dp12 issue d'HS a mis en évidence le rôle de la taille et de la composition des oligosaccharides. Cela a également montré le rôle critique d'un groupement 6-*O*-sulfate qui permet de promouvoir l'activité mitogénique du FGF-2 lors d'interaction avec son récepteur (Pye et al. 1998). Dans ce cas, la longueur de l'oligosaccharide peut aussi jouer un rôle dans l'interaction et l'activation. Enfin, toutes ces caractéristiques sont également capables de moduler la réponse du FGFR avec ses différentes protéines partenaires (Pye et al. 2000).

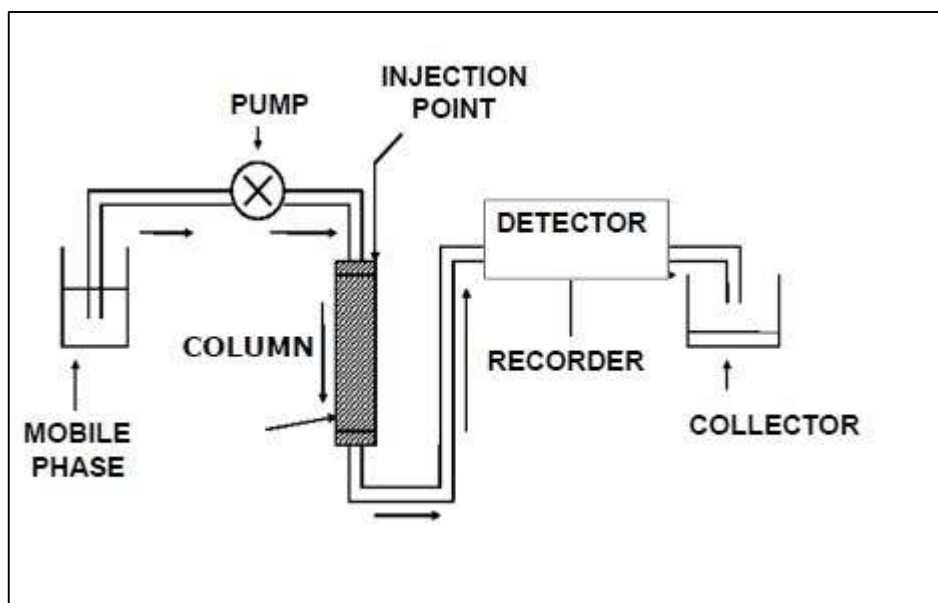


Figure 41 : Schéma d'un appareillage de chromatographie en phase liquide

L'analyse disaccharidique, quant à elle, permet d'analyser et quantifier les modifications (épimérisation et sulfatations) générales des GAG d'un tissu, d'un type cellulaire, ou d'un pôle cellulaire par intégration des différents pics correspondant aux disaccharides. Salmivirta et collaborateurs ont analysé la composition générale des HS au niveau des deux pôles (apical et basolatéral) sur des cellules épithéliales non-différenciées et différenciées issues de la même lignée cellulaire. Ils ont ainsi pu caractériser l'évolution de la structure des HS des cellules Caco-2 (issues de carcinome de colon humain) au cours de leur différenciation. Aucune différence structurale générale n'a été observée entre les deux pôles cellulaires. De même, les domaines *N*-acétylés ne sont pas différents entre cellules non-différenciées et différenciées. Par contre, les régions *N*-sulfatées présentent des variations significatives entre cellules non-différenciées et différenciées. L'épimérisation ainsi que la 2-*O*-sulfatation sont plus importantes lors de la différenciation, alors que la 6-*O*-sulfatation

diminue et que la *N*-sulfatation semble stable (Salmivirta et al. 1998). Toutefois, ce type d'analyse requiert une dépolymérisation complète par digestion enzymatique – qui fait perdre l'information d'épimérisation – ou une dégradation chimique qui ne renseigne pas sur la *N*-sulfatation de la glucosamine. L'une des limitations de la technique est aussi la perte de l'information de séquence des disaccharides au sein des domaines.

La longueur moyenne des chaînes peut également être déterminée par chromatographie, soit par gel filtration *via* la masse molaire relative (M_r) (Trescony et al. 1989), soit par intégration de l'aire des pics correspondant aux disaccharides comparée à celle du tétrasaccharide *linker* insensible aux digestions enzymatiques citées précédemment.

d) Etudes par spectrométrie de masse (MS)

La spectrométrie de masse (MS) est une technique qui permet de mesurer précisément les masses moléculaires. Cette technique nécessite l'ionisation des molécules qui vont ensuite passer à travers un séparateur – généralement un quadripôle ou un temps de vol – avant détection. En fonction de leur rapport m/z (où m est la masse et z leur charge) les espèces ionisées vont arriver au détecteur (Figure 42). La mesure de ce rapport permet de remonter à la masse moléculaire de ces espèces. Appliquée aux oligosaccharides, la mesure de la masse par MS permet de déterminer la longueur ainsi que la composition globale en sulfates et acétyles d'une chaîne de GAG. Toutefois, la MS est une technique destructive qui altère la structure de la chaîne lors de l'ionisation, notamment au niveau des groupements sulfates. Afin de conserver l'intégrité de l'oligomère, des techniques ont été élaborées où la sulfatation des GAG est protégée par un peptide de masse connue riche en résidus basiques (Rhomborg et al. 1998). Cette technique permet de déterminer la masse de l'oligosaccharide jusqu'à une composition tétradécasaccharidique.

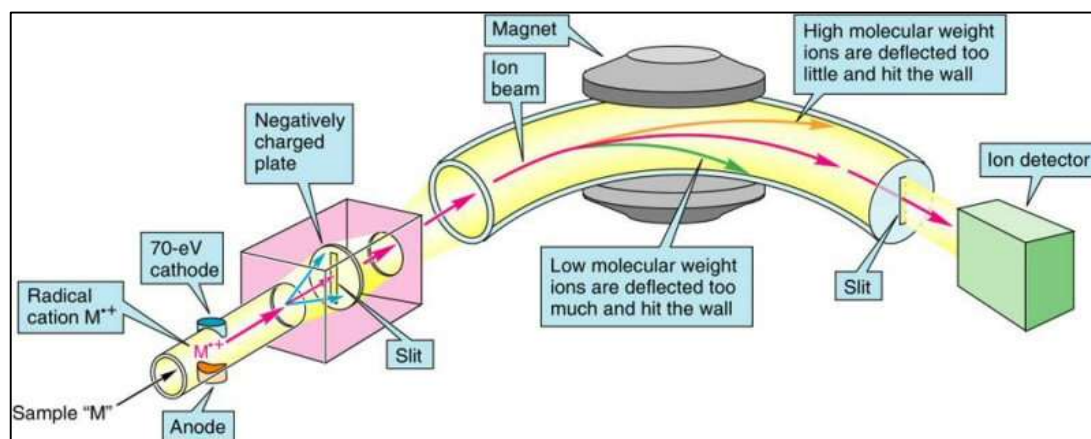


Figure 42 : Schéma d'un appareillage de spectrométrie de masse

La MS peut également être couplée à une technique de purification (HPLC, CE...) pour la détermination de la séquence d'un oligosaccharide inconnu. Comme pour la détermination de séquence par chromatographie, la digestion préalable chimique ou enzymatique des oligosaccharides en disaccharides est nécessaire afin de distinguer différents isomères et leur localisation au sein de l'oligosaccharide (Venkataraman 1999). Parmi les techniques de MS, une technique d'un très grand intérêt pour les séquençages d'oligosaccharides est la

spectrométrie de masse en tandem (ou MS-MS). Après passage par un premier secteur de séparation, les molécules (ou fragments ionisés), d'un ratio m/z unique, sont refragmentées dans une chambre dite de collision puis séparées une nouvelle fois avant détection dans un deuxième secteur de séparation (quadripôle, temps de vol...). Dans le cas des CS, cette méthode a permis de déterminer la position de sulfatations au niveau de la galactosamine et l'épimérisation de l'acide uronique selon les différentes probabilités de former tels ou tels fragments (Lamb et al. 1992) (Zaia et al. 2003). Les différentes techniques de MS bien que destructives requièrent très peu de matériel (à partir d'une centaine de fmoles) et permettent de donner des informations de masse d'une grande précision (jusqu'au dixième de Dalton) ainsi que sur la forme isomérique. La détermination de séquence d'oligosaccharides est également possible, ce qui permet de s'affranchir de la radioactivité, technique coûteuse et dont la lourde réglementation peut être un frein à son utilisation.

Nous avons pu voir les principales techniques de détermination de composition et de séquence de fragments de GAG. Ces méthodes d'analyse et de séparation des différentes espèces permettent d'étudier leurs caractéristiques conformationnelles et d'interaction ainsi que leur dynamique au sein d'échantillons homogènes purifiés et caractérisés.

2. Analyses dynamiques et tridimensionnelles des GAG

Dans cette partie, nous allons voir les différentes techniques d'analyse des propriétés des GAG et de leur conformation spatiale.

a) Etudes par microscopie

La microscopie est une technique peu utilisée dans le domaine de la structure des GAG car elle ne permet pas d'atteindre une résolution à l'échelle atomique. Cependant, les techniques de microscopie photonique et électronique sont utilisées pour la localisation des GAG et plus généralement des PG au sein d'un tissu ou d'une surface cellulaire afin d'étudier leur organisation au niveau des MEC et/ou du glycocalyx (Ripellino et al. 1988) (Figure 43). Des techniques plus récentes de microscopie à force atomique (AFM) ont permis toutefois de visualiser *in vitro* la forme de l'aggrécane ainsi que la disposition et la longueur des chaînes de CS qui y sont attachées, à deux stades de développement différents (Ng et al. 2003) (Figure 44).

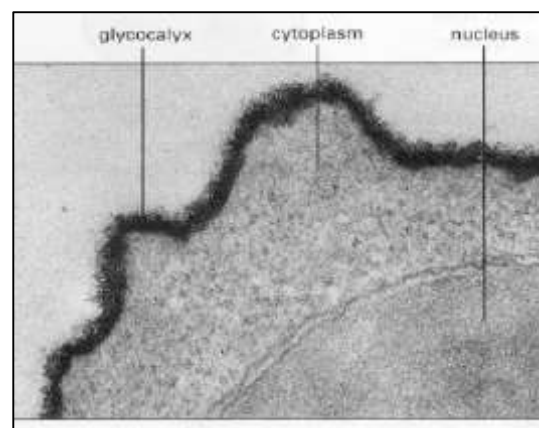


Figure 43 : Cliché d'une coupe transversale de cellule. Le glycocalyx est marqué au rouge de ruthénium. D'après Glauert & Cook dans *Molecular Biology of the Cell*, 4th Edition.

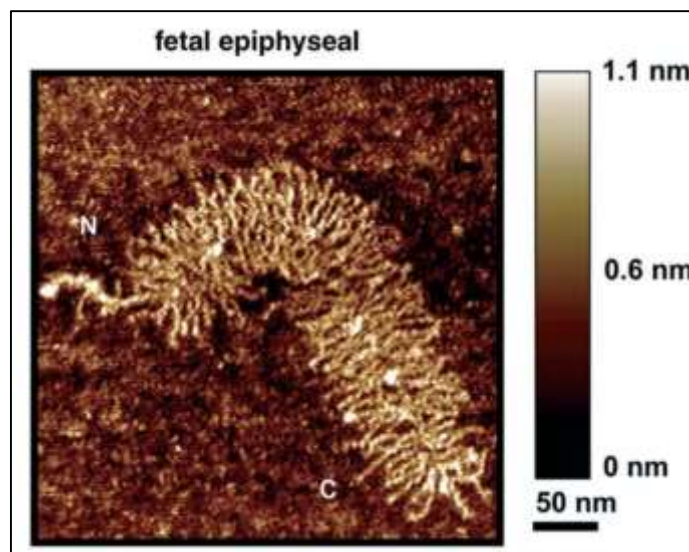


Figure 44 : Image de microscopie à force atomique d'un aggrecan issu d'épiphyse fœtale. D'après (Ng et al. 2003).

b) Etudes par diffraction aux rayons X

Les premières études par diffraction aux rayons X ont été réalisées dans les années 1970 sur des GAG sous forme de fibres oligosaccharidiques. Les clichés suggèrent une forme hélicoïdale des GAG observés dans ces conditions expérimentales avec des pas d'hélice variant selon le GAG et les cations présents dans la structure (Nieduszynski & Atkins 1973) (Atkins & Laurent 1973) (Atkins et al. 1974) (Winter & Arnott 1977) (Winter et al. 1978) (Cael et al. 1978) (Figure 45). Ces structures en hélice présentent une conformation plutôt étendue, stabilisée par les répulsions électrostatiques et des liaisons hydrogènes intramoléculaires (Winter et al. 1978). Cela permet d'avoir une vision structurale au sein des GAG, et notamment au niveau de la conformation des cycles osidiques et des liaisons entre eux. Dans le cas des Hp/HS et DS, différentes conformations thermodynamiquement possibles de l'acide uronique ont été proposées pour rendre compte des résultats des études de diffraction. Il faudra attendre d'autres études réalisées par d'autres méthodes d'analyse comme la RMN pour déterminer plus précisément les différents conformères de l'acide iduronique en solution (Casu et al. 1986). Ces études de diffraction ont permis de déterminer des angles de torsion Φ et Ψ (IUPAC-IUB Joint Commission on Biochemical Nomenclature 1983) préférentiels que peuvent adopter les différentes liaisons des différents résidus saccharidiques (Figure 46). Cela permet de pouvoir comparer les différentes analyses structurales entre elles. Ces données

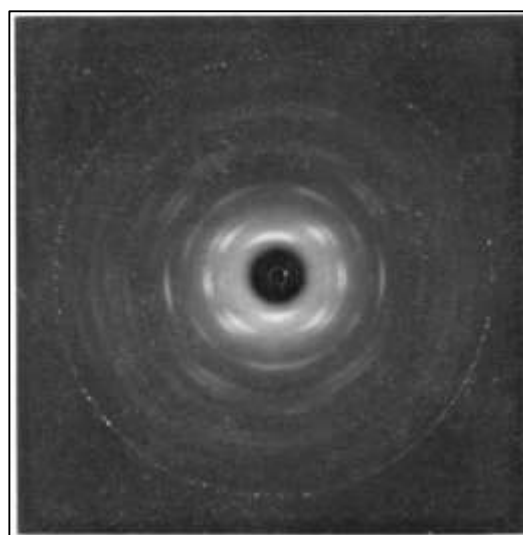


Figure 45 : Premier cliché de diffraction de fibres d'héparine sodique. D'après (Nieduszynski & Atkins 1973).

empiriques ont servi pour les études de modélisation des GAG comme de dynamique moléculaire (Bathe et al. 2005).

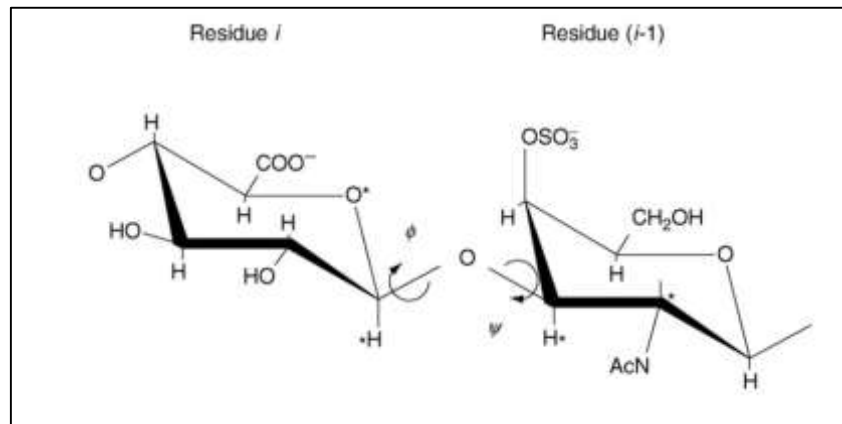


Figure 46 : Nomenclature des angles Φ et Ψ des unités saccharidiques de GAG. Exemple d'un C4S.

Les premières structures de GAG à une résolution atomique ont été obtenues à partir de cristaux d'oligosaccharides complexés à un partenaire protéique (Faham et al. 1996). Les GAG, même à partir de solutions homogènes, possèdent une flexibilité intrinsèque qui ne permet pas un arrangement cristallin. En revanche, lorsqu'un oligosaccharide est en interaction forte avec une protéine, celle-ci bloque ce fragment dans une conformation rigide, et il est possible de cristalliser le complexe ainsi formé. L'un des premiers composés dont la structure tridimensionnelle a pu être ainsi obtenue est l'ATIII en interaction avec un pentasaccharide synthétique d'héparine (Jin et al. 1997). Les résultats de diffraction permettent de reconstruire la structure à une résolution de 2,9Å (0,29nm). Cela a permis de

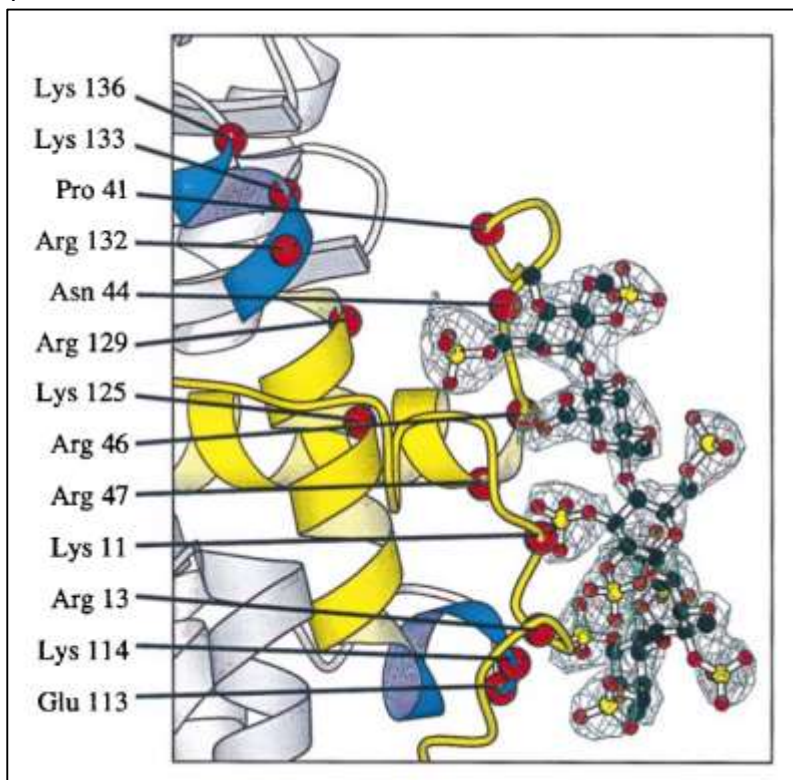


Figure 47 : Structure d'interaction d'un pentasaccharide d'héparine avec l'antithrombine-III. D'après (Jin et al. 1997).

caractériser le mode d'interaction de la séquence d'héparine, en particulier les groupements critiques pour la liaison ainsi que la conformation 2S_0 de l'IdoA,2S (code PDB : 1AZX) (Figure 47). Les études cristallographiques réalisées actuellement portent majoritairement sur des complexes de fragments saccharidiques avec des enzymes de dégradation, souvent bactériennes. Ces études ont trait à la spécificité des enzymes ainsi que leur mode de catalyse

envers des séquences de GAG particulières ou plus larges (Michel et al. 2004) (Shaya et al. 2010).

c) Etudes par diffusion des rayons X aux petits angles (SAXS)

La technique de SAXS permet la détermination de propriétés structurales globales comme la longueur maximale d'un oligosaccharide ainsi que son arrangement dans l'espace. Les études par SAXS des HS et Hp ont été réalisées sur des fragments oligosaccharidiques allant de l'hexasaccharide (degré de polymérisation 6 ou dp6) au dp24 (pour les HS) ou dp36 (pour l'Hp). L'analyse par SAXS permet de déterminer la forme de la molécule ainsi que le rayon de giration (R_G) de la molécule analysée qui est en fait la taille globale de celle-ci. Un facteur prenant en compte le degré de courbure de la molécule noté R_{XS} (pour *radius of cross-section*) est également extrait de l'analyse SAXS. Les résultats de ces études pour l'héparine montrent que le degré de courbure (R_{XS}) est d'autant plus important que le polymère est long. Cela signifie qu'il se courbe, et adopte une conformation spatiale non-linéaire. Ce facteur, combiné au R_G permet de déterminer plus précisément la longueur maximale de l'oligosaccharide. Pour un dp24 d'héparine par exemple, la longueur théorique (L_{th}) maximale est de 11,1nm alors que la longueur expérimentale (L_{exp}) déterminée est de 9,2nm (Figure 48) avec un R_G et R_{XS} de 2,7 et 0,5nm respectivement (Khan et al. 2010).

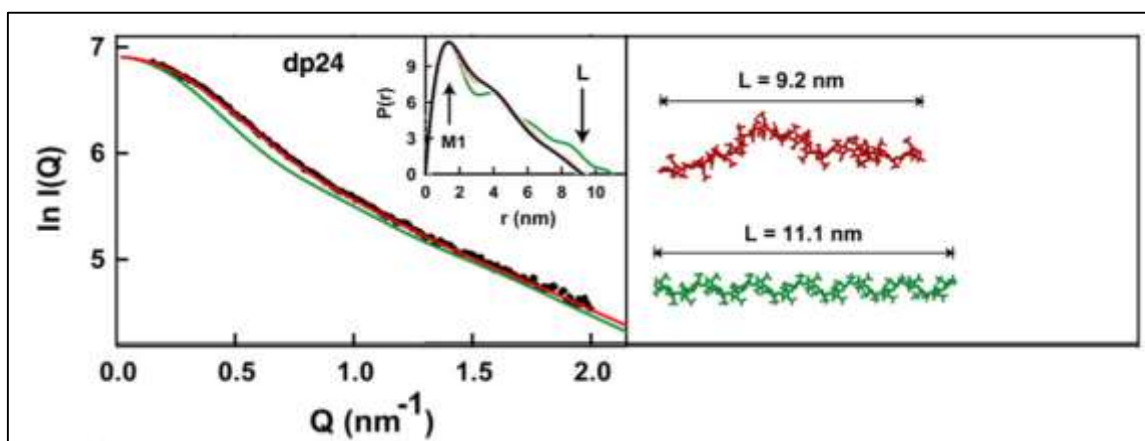


Figure 48 : Courbes d'ajustement de modèles obtenus par diffusion aux rayons-X dans le cas du modèle avec un bon ajustement (rouge) et un mauvais ajustement (vert) du modèle dp24 d'Hp. D'après (Khan et al. 2010).

Dans le cas des HS, les valeurs obtenues expérimentalement pour un dp24 sont les suivantes : $R_G=3\text{nm}$; $R_{XS}=0.65\text{nm}$; $L_{th}=11\text{nm}$; $L_{exp}=8.5-9\text{nm}$. Cela indique que la structure des HS est plus courbée que celle de l'héparine (Figure 49) (Khan et al. 2013). Ces derniers résultats montrent une plus grande flexibilité des HS comparés à l'héparine. Cette flexibilité observée en solution viendrait des différents angles de torsion Φ et Ψ que peuvent adopter les HS comparés à l'Hp, ce qui n'est pas observé à l'état cristallin. Cette apparente flexibilité permettrait aux HS d'interagir plus facilement avec les protéines *in vivo* du fait des multiples orientations que peuvent adopter leurs chaînes.

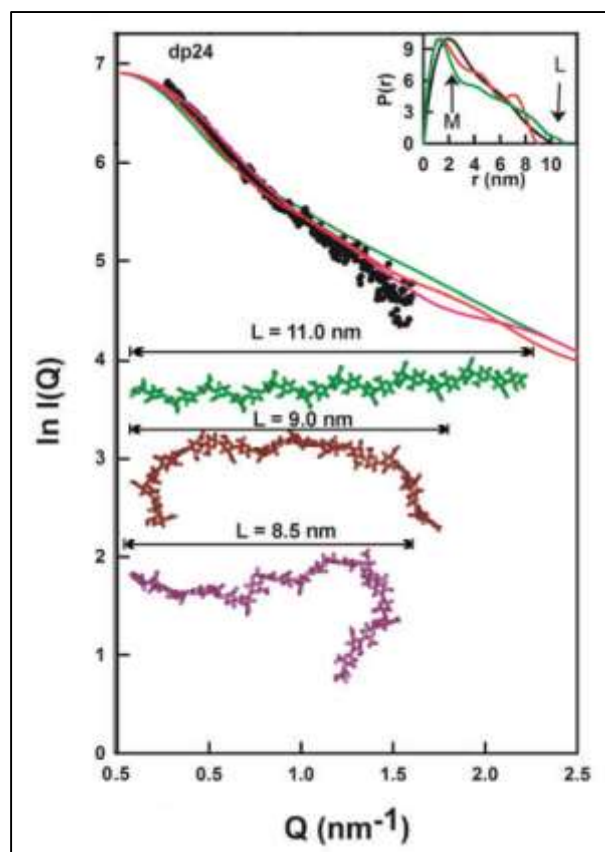


Figure 49 : Courbes d'ajustement de modèles obtenus par diffusion aux rayons-X dans le cas du modèle avec un bon ajustement (rouge et violet) et un mauvais ajustement (vert) du modèle dp24 d'HS. D'après (Khan et al. 2013).

d) Etudes de la dynamique moléculaire (MD)

La dynamique moléculaire est une technique computationnelle qui permet de simuler au cours du temps le comportement d'une molécule évoluant selon les lois physiques (appelées champs de force) imposées par l'expérimentateur. D'autres contraintes sont également déterminées empiriquement à partir des données expérimentales existantes obtenues par des techniques d'analyse spectroscopique, de diffraction *etc.*

Concernant les GAG, les études de dynamique moléculaire ont été focalisées principalement sur des oligosaccharides de taille limitée. De nos jours avec l'amélioration des systèmes d'analyse, ces études sont réalisées sur de plus gros fragments atteignant des degrés de polymérisation de 36 monosaccharides (ou dp36) pour l'Hp (Khan et al. 2010), dp24 pour les HS (Khan et al. 2013), et dp50 pour l'HA (Almond et al. 2006). Les informations d'entrée proviennent des techniques citées précédemment ainsi que de diffusion aux neutrons (Jasnin et al. 2010) et majoritairement des données obtenues par résonance magnétique nucléaire (RMN) (Mulloy et al. 1993) (Mulloy & Forster 2000). Ces informations sont intégrées à un champ de force qui prend également en compte les encombrements stériques, les influences électrostatiques, les angles de torsion, les conformations possibles des cycles osidiques, de l'environnement moléculaire (eau, vide...) *etc.* Les champs de force les plus utilisés sont CHARMM (*Chemistry at HARvard Macromolecular Mechanics*) (Brooks et al. 2009), AMBER

(*Assisted Model Building with Energy Refinement*) (Homans 1990), MM2/MM2CARB (*Molecular Mechanics 2/for CARBohydrate*) (Tvaroška & Pérez 1986). D'autres champs de force ont été adaptés à partir de ceux existants pour les protéines ou nucléotides ou bien créés exclusivement pour les hydrates de carbone (Imberty & Pérez 2000) du fait de l'intérêt croissant qu'ils représentent et de leurs structures particulières. Brièvement, un calcul MD permet la génération d'un grand nombre de structures qui sont ensuite triées en fonction de leur énergie conformationnelle, les structures de plus basse énergie étant les plus stables. L'analyse de ces structures permet de caractériser des conformations spatiales communes et/ou divergentes. Des différences ont ainsi pu être mises en évidence entre les HS et les Hp, les HS pouvant adopter des formes étendues ou courbées alors que les Hp, étant plus rigides, ont une forme principalement étendue *in vitro* (Figures 48, 49).

Les analyses de dynamique moléculaire s'intéressent également de plus en plus aux interactions oligosaccharide-protéine(s). Ces analyses peuvent apporter des informations structurales lorsque les études de ces complexes par voie expérimentale ne sont pas disponibles. Différentes interactions HS/Hp-chimiokines ont été étudiées au laboratoire afin de déterminer le mode de liaison de ces GAG à leur partenaire. A l'aide de programmes déterminant les aires électropositives (comme MOLCAD, GRID) susceptibles d'être en interaction avec les oligosaccharides sulfatés, Lortat *et al.* ont pu déterminer des résidus basiques des protéines jouant un rôle important dans la liaison avec des GAG ainsi que l'orientation de ces derniers. Différents modes d'interaction ont pu être proposés selon la famille de la chimiokine concernée (basée sur la conservation de cystéines proches au sein de la séquence primaire de la protéine : CC, CXC, CX₂C, CX₃C) (Lortat-Jacob *et al.* 2002). D'autres interactions ont également été caractérisées par la suite (Ricard-Blum *et al.* 2004) (Sapay *et al.* 2011) (Laguri *et al.* 2011) (pour plus d'exemples voir la revue (Imberty *et al.* 2007)).

Les analyses des GAG par dynamique moléculaire grâce aux données obtenues par différentes méthodes structurales permettent d'appréhender les GAG et leur comportement, seuls ou lors d'interaction avec un partenaire protéique. Ces dernières analyses permettent de déterminer simplement les résidus d'interaction importants ou nécessaires et peuvent être confirmées par mutagenèse dirigée (Ricard-Blum *et al.* 2004). Des techniques supplémentaires d'amarrage moléculaire (de l'anglais *molecular docking*) permettent de modéliser des structures d'interaction lorsque celles-ci sont difficiles à réaliser avec les techniques actuelles et leurs limitations (taille pour la RMN, cristallogénèse et flexibilité pour leur diffraction aux rayons-X...).

3. La RMN, une technique de choix

a) Principes de la RMN

Ce travail n'ayant pas nécessité de développement de méthodes spectroscopiques pour sa réalisation, des rappels succincts des principes de la résonance magnétique nucléaire seront fait dans cette partie.

Certains noyaux atomiques, en particulier les noyaux d'hydrogène (abrégé en ^1H ou proton), de l'isotope 13 du carbone (noté ^{13}C) et de l'isotope 15 de l'azote (^{15}N) pour les noyaux les plus utilisés en biologie structurale, possèdent une propriété quantique appelée spin nucléaire dont l'image la plus simple est celle d'un petit aimant. Ces noyaux sont caractérisés par un spin nucléaire (I) et un rapport gyromagnétique (γ) non nul qui est propre à chaque type de noyau. Placés dans un champ magnétique, ils vont avoir tendance à s'orienter suivant celui-ci (selon un axe arbitraire z d'un repère orthogonal) parallèlement ou antiparallèlement (Figure 50). Lorsqu'une impulsion de radiofréquence (RF) est appliquée au système et que celle-ci a la même énergie que l'énergie séparant les états parallèles et antiparallèles (la séparation en énergie étant proportionnelle au rapport gyromagnétique et au champ), les spins vont transiter entre les deux états ce qui est le phénomène de résonance. En fait, cette description quantique s'applique aisément lorsque l'on considère des spins isolés

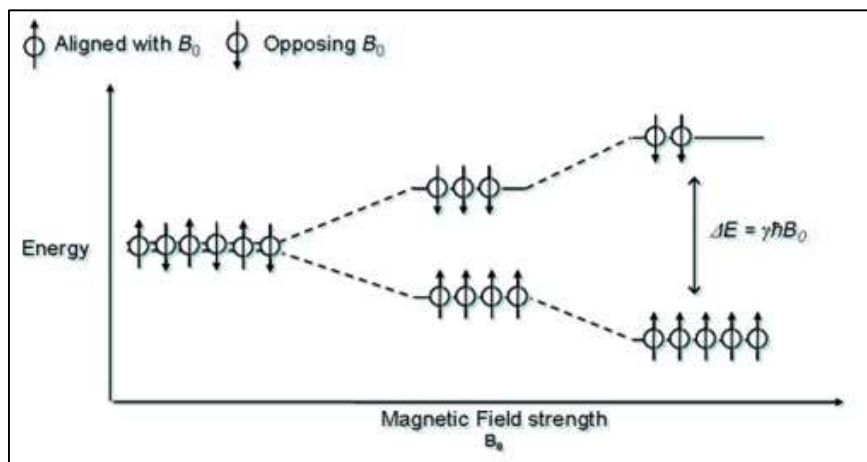


Figure 50 : Représentation de spins nucléaires en fonction de la force du champ magnétique dans lequel ils sont placés.

mais elle devient difficile à mettre en œuvre pour une population de spins et il est utilisé une description en aimantation globale suivant les équations classiques de Bloch.

Dans cette description, les spins ont des orientations aléatoires en l'absence de champ mais s'alignent, lorsqu'ils sont placés dans un champ magnétique intense (B_0), dans la direction du champ en tournant autour de lui avec une fréquence de rotation dépendant en première approximation du champ et du rapport gyromagnétique. L'application de ce champ crée une différence de population entre les états excités et fondamentaux ce qui se traduit par l'apparition d'une aimantation globale non nulle. Lorsqu'une radiofréquence (RF) est

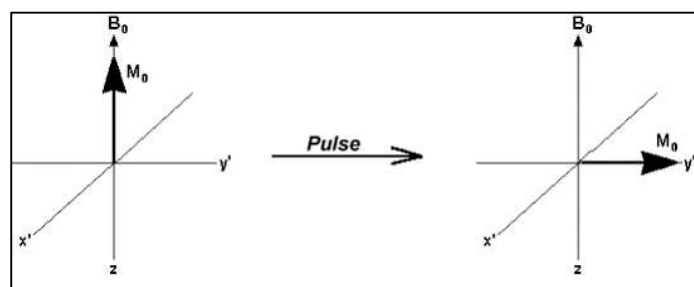


Figure 51 : Représentation d'une aimantation après une impulsion radiofréquence

appliquée au système et que celui-ci a la même fréquence que la fréquence de rotation des spins, l'aimantation de ces spins dévie de son axe selon un certain angle qui dépend de la durée (τ) de l'impulsion RF (Figure 51). Suite à l'extinction du signal RF, le système va revenir à l'équilibre et va réémettre lors de ce retour un signal de radiofréquence qui pourra alors être détecté par une bobine du spectromètre (Figure 52). En fait, chaque spin est dans un

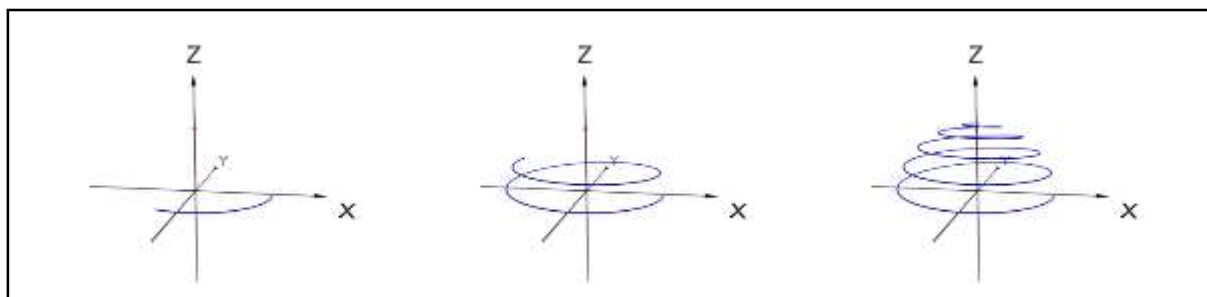


Figure 52 : Schéma de la relaxation de spin au cours du temps.

environnement chimique qui lui est propre et qui crée autour de lui un petit champ local caractéristique. Les fréquences de précession des différents spins ne sont donc pas identiques. Leur excitation est faite, dans les techniques RMN utilisées maintenant, par une impulsion de courte durée et de forte puissance qui affecte l'ensemble des spins d'un même type (^1H , ^{13}C ou ^{15}N par exemple) et le signal détecté ou FID (pour *Free Induction Decay*) est donc composé d'une somme de signaux d'amplitudes et de fréquences différentes. La déconvolution de ce signal par Transformée de Fourier (FT) permet de séparer les différentes fréquences qui le composent et donc de différencier les signaux correspondant aux différents noyaux (Figure 53).

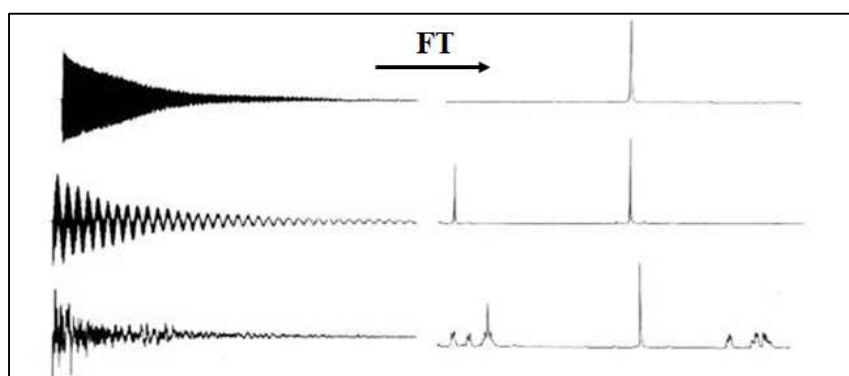


Figure 53 : Transformée de Fourier d'un signal FID d'un (haut), deux (milieu) ou plusieurs (bas) spins de noyaux chimiquement non-équivalents.

L'analyse de ces signaux permet d'obtenir une mesure des observables du système : intensité, déplacements chimiques, couplages scalaires, couplages dipolaires et relaxation dont de brefs rappels des origines physiques vont être faits ainsi que des informations qu'il est possible d'en obtenir.

L'intensité du signal de RMN dépend de l'aimantation nette induite par le champ magnétique. Cette aimantation reste faible malgré les champs appliqués qui peuvent atteindre jusqu'à 20 Teslas – la différence de population entre états fondamentaux et excités étant de l'ordre du 1/10000. C'est pourquoi la RMN est une technique peu sensible, et elle

nécessite une quantité importante de matériel – en solution des concentrations de l'ordre de 100 micromolaires à plusieurs millimolaires. Afin d'améliorer le rapport signal/bruit, il est possible d'augmenter le nombre d'itération de l'expérience RMN comme d'utiliser des sondes opérant à très basse température pour réduire le bruit thermique. Enfin, pour l'hydrogène, la proportion de l'isotope ^1H observable est de 99,99%. Cela n'est pas le cas pour tous les atomes et l'isotope observable par RMN n'est pas forcément le plus abondant. Par exemple, pour les deux isotopes les plus fréquemment employés en biologie structurale, le ^{13}C et ^{15}N , l'abondance naturelle est de 1,1% et 0,36% respectivement (Figure 54). Il est donc nécessaire

Nucleus	Spin (I)	γ rad T ⁻¹ s ⁻¹	Natural abundance
^1H	1/2	2,675x10-8	99,98 %
^2H	1	0,411x10-8	0,016 %
^{31}P	1/2	1,083x10-8	100,00 %
^{15}N	1/2	-0,271x10-8	0,365 %
^{13}C	1/2	0,673x10-8	1,11 %
^{19}F	1/2	2,627x10-8	100,00 %

Figure 54 : Tableau regroupant les principaux isotopes observables par RMN, leur nombre de spin, leur rapport gyromagnétique ainsi que leur abondance naturelle.

pour accroître la sensibilité d'enrichir les échantillons en ces isotopes de façon globale ou spécifique en fonction des informations souhaitées.

Le déplacement chimique (delta). Chaque noyau est sous l'effet du champ magnétique intense B_0 mais également d'un champ local induit par l'environnement électronique (ou chimique). Ce dernier est généralement différent pour deux noyaux en dehors de centres de symétrie. Les fréquences de résonance des différents noyaux constituent les déplacements chimiques et sont donc dépendants du champ local et de B_0 . Afin de pouvoir comparer plusieurs échantillons étudiés dans des spectromètres de champs différents, les fréquences observées sont comparées à une fréquence de résonance connue (ν_{ref}) par rapport à la fréquence de travail (ν_0 ou ν_{travail}). La valeur est exprimée en parties par millions (ppm) et est indépendante de l'intensité du champ magnétique du spectromètre.

Le couplage scalaire en RMN est l'interaction entre spins nucléaires *via* les liaisons chimiques. Ce couplage provoque un éclatement du signal RMN, c'est-à-dire une multiplicité des pics observés. Cet éclatement du signal dû à l'interaction d'un (ou plusieurs) voisin est riche en informations. En plus de renseigner sur le nombre de spins qui sont couplés au noyau considéré, la valeur (en Hz) de cet écart des signaux, appelée constante de couplage et noté $^nJ_{ij}$ (où i et j sont les noyaux considérés et n le nombre de liaisons chimiques séparant les deux noyaux), contient des informations angulaires (Haasnoot et al. 1980) et donc structurales. En biologie structurale, les techniques de découplage sont généralement utilisées afin d'éviter cette augmentation du nombre de pics, et donc de simplifier l'analyse spectrale.

Le couplage dipolaire, contrairement au couplage scalaire, est l'interaction entre spins à travers l'espace. Ces couplages sont moyennés à zéro en solution à cause des mouvements

moléculaires et donc ne se traduisent pas par un effet directement observable sur les spectres. Ils sont toutefois responsables de transferts de polarisation d'un spin à un autre dont l'intensité dépend de la distance entre ces spins. Cet effet a une grande importance dans les études structurales puisqu'il permet de déterminer les atomes proches voisins dans l'espace. En particulier, des expériences de type NOESY (*Nuclear Overhauser Effect Spectroscopy*) permettent d'obtenir ces informations de distance qui peuvent être calculées à partir de l'intensité des signaux grâce à la formule Figure 55 (Wüthrich 1986). Cependant le signal obtenu permet d'obtenir une information de la distance moyenne entre les deux spins, la molécule n'étant pas rigide, l'espace entre les spins peut varier en fonction des mouvements moléculaires. Malgré cette ambiguïté, des informations de fortes, moyennes, ou faibles interactions peuvent être déterminées et être utilisées pour les calculs de structures tridimensionnelles de la molécule et de dynamique moléculaire.

$$nOe_{\text{intensity}} \propto \frac{1}{r^6} \cdot f(\tau_c)$$

Figure 55 : Facteurs proportionnels à l'intensité de signaux nOe. r^6 : distance entre deux noyaux. τ_c : temps de corrélation.

Le temps de corrélation (τ_c) est enfin, un autre paramètre moléculaire à prendre en compte dans le cas de la RMN en solution. Il représente le temps moyen pour une molécule de faire une rotation d'un radian. Ce τ_c est dépendant de la taille de la molécule, de la température et de la viscosité du milieu. Plus l'objet considéré a une taille importante et plus le milieu est visqueux, plus long sera le τ_c (Figure 56). Plus le temps de corrélation est long,

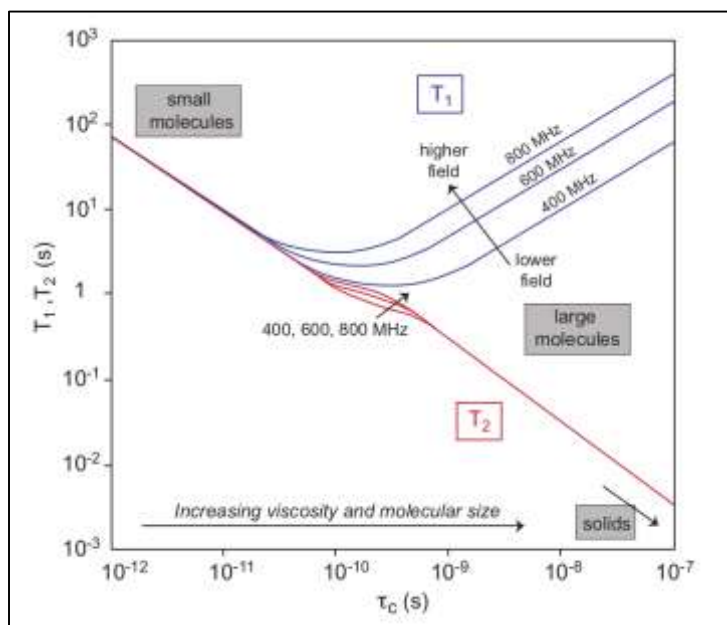


Figure 56 : Evolution des temps de relaxation T_1 et T_2 en fonction du temps de corrélation (τ_c) et de l'intensité du champ magnétique.

plus le temps de relaxation transversale (appelé T_2) est court. Or, le T_2 est relié à la largeur du signal du noyau considéré, et la largeur du pic à mi-hauteur est égale à $1/\pi T_2$ (Figure 57). En résumé, plus la molécule est de taille importante, plus sa rotation d'un radian va être longue

(τ_c important) donc plus le temps de relaxation transversale va être court, donc plus les signaux seront larges. L'élargissement des signaux peut être tel pour les grosses molécules que les signaux se confondent entre eux ou avec le bruit de fond, et deviennent indétectables. La RMN est donc limitée pour l'étude des objets de grandes tailles (supérieurs à environ 40kDa) sans des techniques de marquage spécifiques. Une façon d'améliorer la qualité/résolution des spectres est donc de travailler à une température la plus haute possible mais restant compatible avec l'échantillon et le matériel (afin d'augmenter l'agitation thermique et de réduire le τ_c).

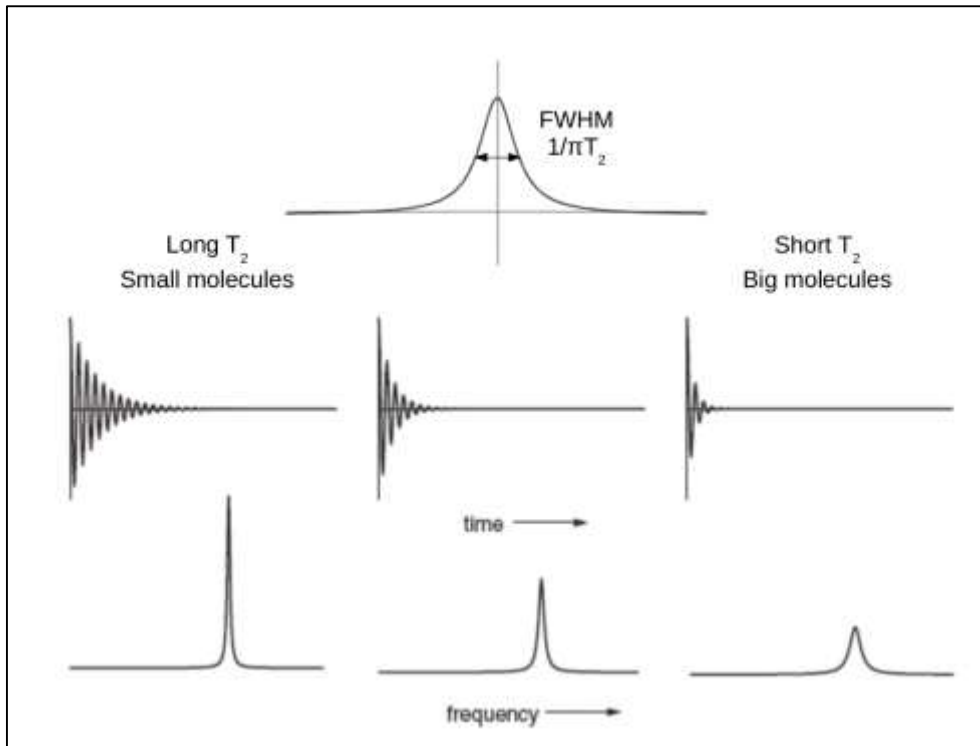


Figure 57 : Effets du temps de relaxation transversale T_2 sur le signal FID brut (milieu) et après transformée de Fourier (bas). La valeur de la largeur à mi-hauteur (full width at half maximum, FWHM) est dépendante du T_2 (haut).

Les temps de relaxation, en RMN, sont le retour des spins d'un état excité à un état d'équilibre. Deux types de relaxations sont considérés :

* La relaxation longitudinale T_1 . C'est le temps que mettent les spins excités à revenir selon l'axe z parallèle au champ B_0 . Elle correspond au retour de l'aimantation maximale du système et définit le temps d'attente nécessaire entre deux séquences impulsionnelles (Figure 58).

$$M_z(t) = M_0(1 - 2e^{-t/T_1})$$

Figure 58 : Equation de détermination de l'aimantation selon l'axe z au temps « t » $M_z(t)$ en fonction de l'aimantation initiale M_0 et du temps de relaxation longitudinale T_1 .

* La relaxation transversale T_2 . C'est le temps de disparition par défocalisation de l'aimantation selon l'axe xy perpendiculaire à B_0 . C'est cette relaxation qui limite la détection du signal RMN et qui est dépendante du τ_c (Figure 59).

$$M_{x,y}(t) = M_0 \cdot e^{-t/T_2}$$

Figure 59 : Equation de détermination de l'aimantation selon l'axe x,y au temps « t » $M_{x,y}(t)$ en fonction de l'aimantation initiale M_0 et du temps de relaxation transversale T_2 .

Les mécanismes de relaxation sont reliés aux mouvements internes de la molécule mais également aux mouvements globaux de celle-ci en solution (et donc au τ_c). C'est pourquoi les temps de relaxation (T_1 et T_2) ou vitesses de relaxation (R_1 et R_2 , respectivement l'inverse de T_1 et T_2) sont utilisés pour observer la dynamique des différents résidus et le comportement de la molécule en solution.

b) Etudes de la structure et de la dynamique des GAG par RMN en solution

Les glycosaminoglycanes sont des molécules extrêmement solubles et donc leurs études se sont faites très majoritairement en solution. Les principales caractéristiques saccharidiques (composition, longueur), conformationnelles et spatiales des oligosaccharides ainsi que leur dynamique ont été déterminés à partir des différentes observables citées précédemment.

-Analyse des déplacements chimiques : l'attribution des signaux des spectres RMN de GAG a débuté par l'analyse de disaccharides et oligosaccharides connus. Des méthodes en amont ont été utilisées pour digérer, purifier et caractériser les différentes unités disaccharidiques. Celles-ci ont ensuite été analysées par RMN monodimensionnelle (^1H et ^{13}C) dans un premier temps (Huckerby & Nieduszynski 1982) (Casu et al. 1981) et l'attribution a été réalisée grâce à des techniques de découplages nucléaires et des différents déplacements chimiques de composés soumis à des digestions enzymatiques ou dégradations chimiques spécifiques (Yates et al. 1996). Une grande quantité de déplacements chimiques des noyaux de fragments saccharidiques a été déterminée depuis pour les HS, Hp et structures analogues, CS/DS, KS et HA (pour revue voir (Pomin 2013) (Figure 60).

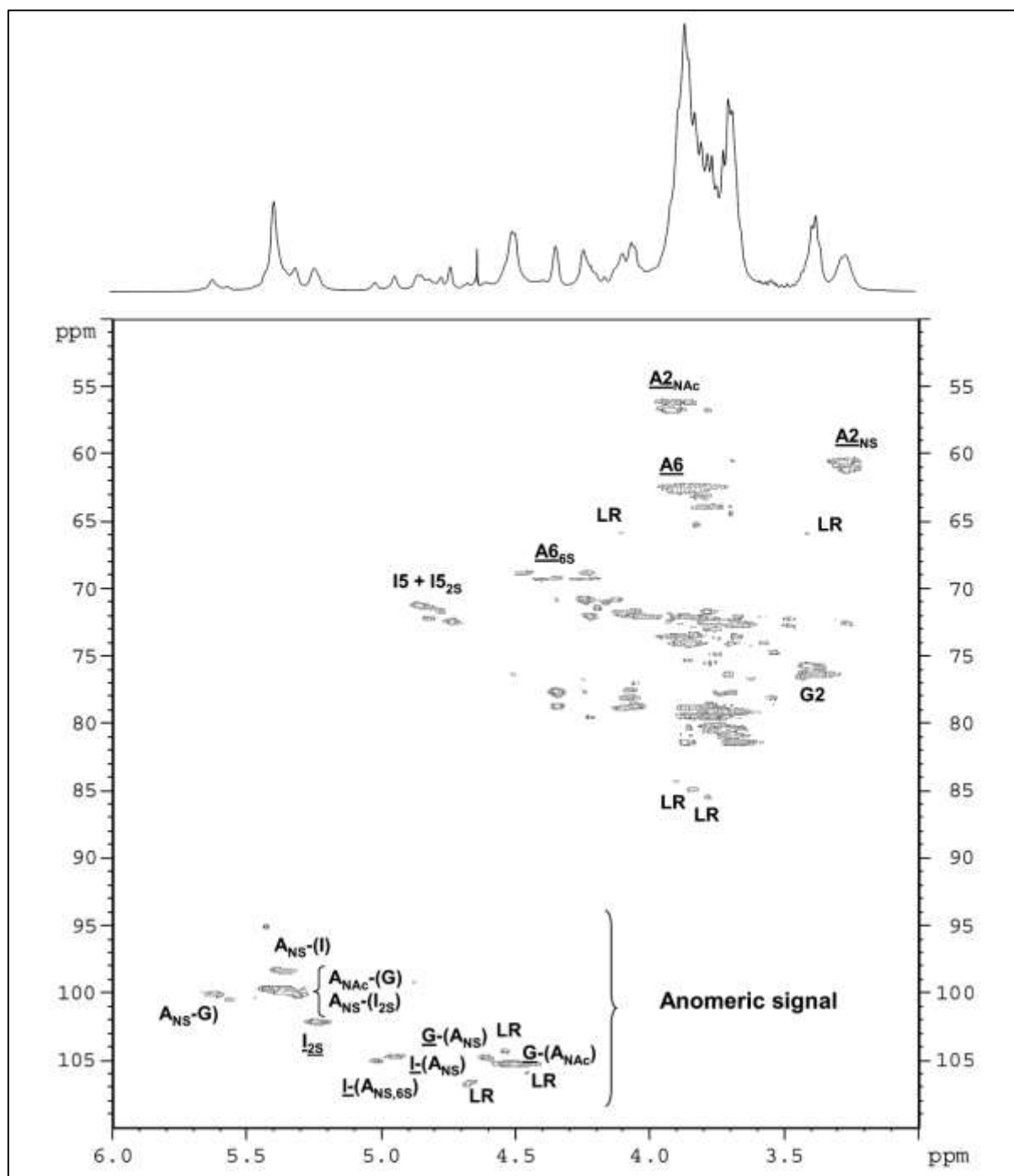


Figure 60 : Spectres ^1H et ^1H - ^{13}C -HSQC d'HS non-enrichi en isotopes. G : acide glucuronique ; I : acide iduronique ; A : glucosamine ; LR : région de liaison (tétrasaccharide linker). D'après (Guerrini et al. 2005).

-Intégration de l'intensité des signaux observés : l'intégration des intensités des pics obtenus par RMN du proton, directement corrélées à la quantité de l'espèce observée, permet de déterminer le pourcentage des différentes modifications en calculant le rapport entre les diverses espèces présentes dans le signal RMN. Cette quantification peut être également obtenue par des techniques de RMN bidimensionnelles en considérant le noyau du ^{13}C (Guerrini et al. 2001). Avec les déplacements chimiques, ces deux données permettent de caractériser la composition et le degré de pureté d'un échantillon, ce qui est primordial pour le contrôle qualité des GAG utilisés dans le domaine de la Santé. L'efficacité de la RMN a été démontrée pour différencier des lots d'Hp purs et des lots contaminés à l'OSCS (Liu et al.

2009). La longueur moyenne de GAG entiers peut également être déterminée par rapport à l'intensité du signal du xylose (Iacomini et al. 1999) présent dans le tétrasaccharide *linker* (à l'exception des HA) comparée à l'intensité des signaux du motif disaccharidique.

-Etude des constantes de couplages : dans les GAG, les valeurs des constantes de couplage permettent de déterminer la position axiale ou équatoriale des groupements substitués, la stéréoconfiguration de l'ose (α/β) et la conformation de l'acide iduronique (Tvaroška et al. 1989) (Bose et al. 1998) (Ferro et al. 1990) (Casu et al. 1986). Ainsi, dans l'héparine, il a pu être montré l'influence sur la conformation 1C_4 , 2S_0 et 4C_1 de l'iduronate (Figure 61) des substitutions dudit iduronate et de celles des oses adjacents (Murphy et al. 2008) comme des contrions présents en solution (Hricovini 2011).

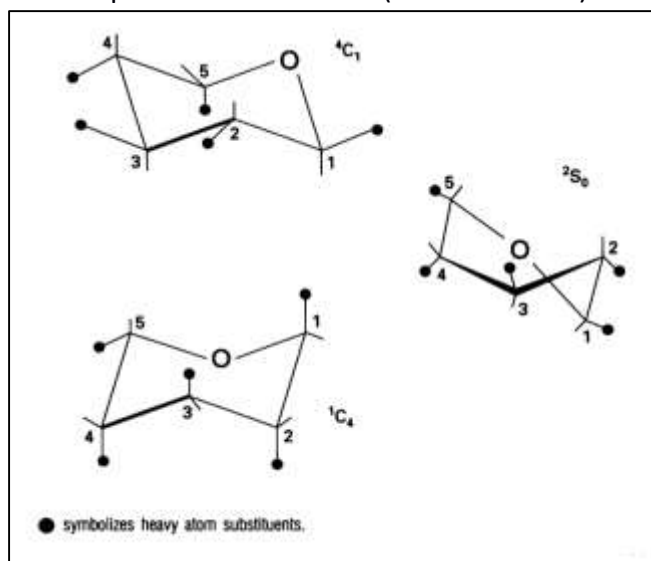


Figure 61 : Trois conformations principales de basse énergie de l' α -L-idopyranose retrouvé en solution. D'après (Mulloy et al. 1993).

-Les effets Overhauser nucléaires (nOe) : les nOe sont relatifs aux distances inter-atomiques, et ont permis pour des fragments d'héparine de déterminer leurs possibles structures en solution en fonction des différentes modifications oligosaccharidiques ainsi que de prédire certaines interactions (Mulloy et al. 1993) (Mikhailov et al. 1996) (Mulloy & Forster 2000) (Figure 62). Ces études ont également montré l'influence des O-sulfatations sur la géométrie globale de l'héparine plus que la C5-épimérisation et que la N-substitution (Mulloy et al. 1994) (Zhang et al. 2008). La conformation du cycle de l'acide uronique par étude des nOe a montré l'influence des contrions Ca^{2+} comme favorisant fortement la forme 1C_4 de l'IdoA,2OH par rapport aux contrions Na^+ , K^+ et Mg^{2+} (Rudd et al. 2007) et peut donc influencer la fixation aux protéines et leurs activités. Différents rapports entre les populations des conformères de l'acide iduronique ont été déterminés par Mulloy *et al* (Mulloy et al. 1993) (Mulloy et al. 1994) et par Cros *et al* 1997 (Cros et al. 1997) pour des tétrasaccharides et trisaccharides respectivement. La taille de la molécule aurait donc également un rôle dans la structure du sucre.

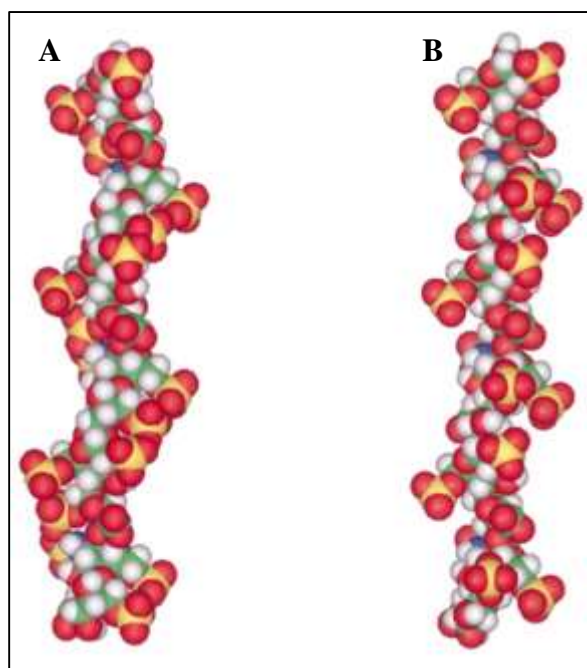


Figure 62 : Structure d'héparine en solution déterminée par RMN selon la conformation des iduronates en 1C_4 (A) ou 2S_0 (B). D'après (Mulloy & Forster 2000).

-Les temps de relaxation : les temps de relaxation des GAG sont également étudiés par RMN. Une étude de la dynamique d'un pentasaccharide synthétique d'héparine a montré que les temps de relaxation T_1 varient suivant la position des protons au sein de la chaîne comme à l'intérieur même d'un résidu osidique (Hricovlni & Torri 1995). Ces résultats indiquent un mouvement anisotrope de la molécule. Les temps de relaxation T_1 sont également différents selon la position des résidus au sein de la chaîne ; les résidus des côtés non-réducteur et réducteur étant plus libres que ceux qui, à l'intérieur de la chaîne, sont contraints par deux liaisons osidiques. Cette plus grande flexibilité des résidus au niveau des extrémités est observée dans les oligosaccharides d'HS issus des domaines NA (Hricovini et al. 1997) et d'HA (Cowman et al. 2001). La flexibilité dépend également de la présence de résidus IdoA, ces derniers pouvant adopter trois différentes conformations spatiales (1C_4 , 2S_0 et 4C_1). La flexibilité est également différente selon le motif de sulfatation (Angulo et al. 2005) et les contrions présents en solution, les cations divalents (Ca^{2+} et Mg^{2+}) introduisant plus de rigidité au sein de la chaîne (Rudd et al. 2007).

c) Enrichissement isotopique des GAG

L'utilisation dans des échantillons naturels de techniques RMN multidimensionnelles hétéronucléaires (1H - ^{13}C et 1H - ^{15}N) a été reportée dans un certain nombre d'études. Toutefois, cette utilisation est restée restreinte en raison des quantités importantes d'échantillon (de l'ordre de 50 à 100mg) que nécessitent ces techniques. Bien qu'il serait certainement d'un très grand intérêt pour un grand nombre d'études RMN de disposer de GAG marqués, il n'y a eu jusqu'à maintenant que très peu de synthèse de ces composés.

La majorité des productions de chaînes d'oligosaccharides marqués ^{13}C ou ^{15}N ont été faites à partir de la souche d'*E.coli* K5. Cette souche a la particularité de produire une capsule polysaccharidique dont la structure primaire (GlcNAc-GlcA)_n appelée héparosan est analogue aux HS non-modifiés. La culture de cette souche dans un milieu comprenant du ^{13}C -glucose et/ou du ^{15}N -ammonium conduit à la biosynthèse de ce polymère marqué et peut être utilisé dans des expériences d'interaction après modifications chimio-enzymatiques du polysaccharide (Laguri et al. 2011). En 2004 le groupe d'Almond qui, en introduisant dans cette souche bactérienne le gène de la HAS, a pu obtenir des échantillons d'acide hyaluronique diversement marqués. Cela a permis d'attribuer complètement les signaux des unités GlcN de chaînes jusqu'à 10 unités saccharidiques (Blundell et al. 2004). Cette technique de marquage ^{15}N a par la suite été utilisée pour déterminer les paramètres dynamiques et conformationnels des chaînes d'hyaluronate (Almond et al. 2006). Pour les unités glucuroniques GlcA, l'attribution s'est révélée plus difficile malgré un marquage au ^{13}C et s'est limitée à des séquences de tétra- et hexasaccharides (Colebrooke et al. 2005).

Il est également possible sans marquer l'ensemble de la chaîne, d'introduire localement des sondes marquées. Ainsi, dans des chaînes de chondroïtines sulfates une désacétylation complète des unités GlcN suivie de leur réacétylation avec des acétyles marqués au ^{13}C a permis de marquer un pentasaccharide et, en conjonction avec d'autres données structurales, d'en déterminer la structure (Yu et al. 2007).

L'utilisation de marquage isotopique ^{13}C et ^{15}N constitue donc une grande avancée sur le plan de la structure, de l'interaction et de la dynamique des GAG. Par contre, cette voie n'a été que très peu explorée dans la caractérisation de la structure et de l'organisation des GAG à la surface cellulaire. Il n'existe qu'une étude faisant appel à un marquage isotopique pour la caractérisation de ce type de GAG. Pour cela, des cellules endothéliales de souris ou des cellules d'ovaire de hamster chinois (CHO) ont été cultivées en présence de ^{15}N -glutamine (Figure 63). Après lyse des cellules, puis purification des GAG, la comparaison avec des spectres RMN d'échantillons d'HS, CS ou Hp à l'abondance naturelle les spectres montrent essentiellement la présence des signaux des CS pour les cellules épithéliales (Pomin et al. 2010). A l'inverse, les cellules CHO présentent des signaux attribuables majoritairement aux azotes de la fonction amide *N*-acétyl des GlcNAc. La quantification de ces signaux montre ainsi une répartition 75:25 entre les enchaînements GlcNAc-GlcA et GlcNAc-IdoA dans ces GAG cellulaires. Toutefois, cette méthode reste limitée pour l'analyse des chaînes d'HS. Le nombre de sondes disponibles dans la chaîne est restreint puisqu'il n'y a qu'une fonction amide par disaccharide. Ensuite, dû aux échanges des protons avec le solvant, les fonctions *N*-sulfatées ne sont pas visibles ce qui réduit encore le nombre de positions observables. Enfin, le taux de marquage des fonctions amides par ce protocole reste faible, de l'ordre de 10% ce qui limite fortement la sensibilité.

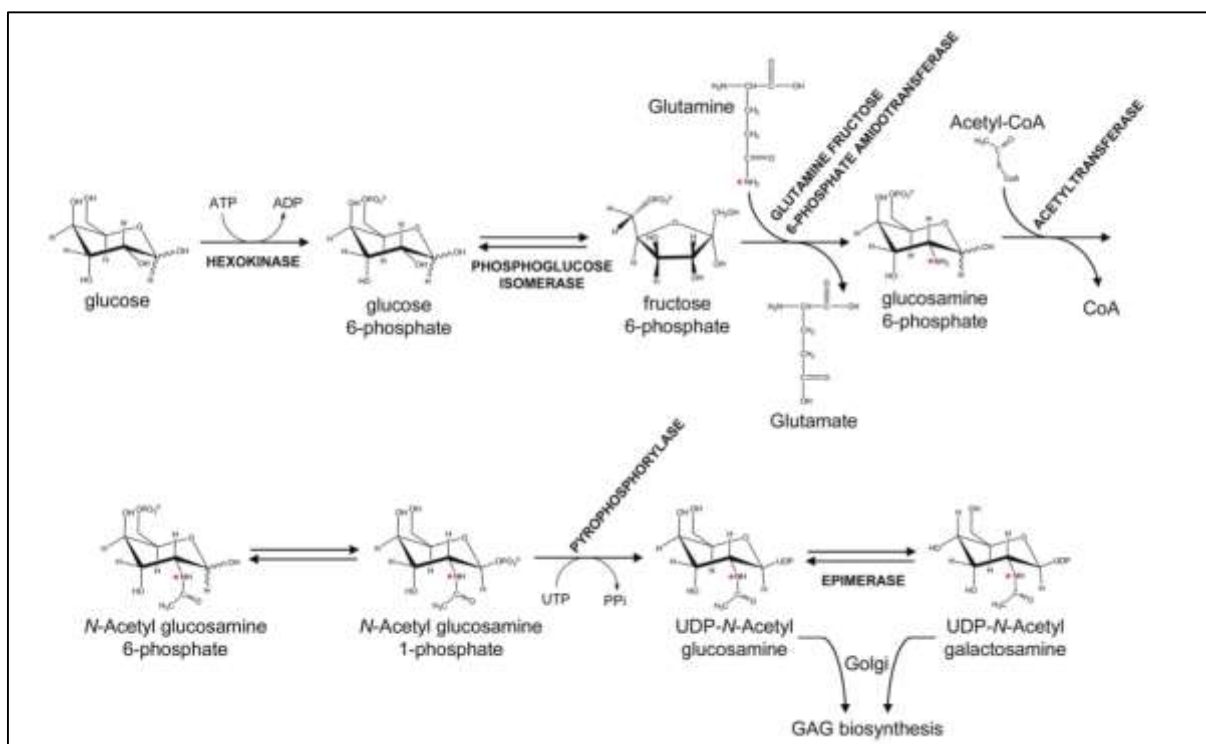


Figure 63 : Voie métabolique de la biosynthèse d'UDP-GlcNAc et d'UDP-GalNAc marqués à l'isotope 15 de l'azote à partir d'un précurseur de ¹⁵N-glutamine. D'après (Pomin et al. 2010).

d) Etudes d'interaction HS-protéine par RMN

Les interactions des HS/Hp avec leur partenaire protéique ont été intensivement étudiés par RMN (voir (Pomin 2014) pour revue sur les travaux récents). Ces études RMN nécessitent l'attribution des signaux de résonances des différents résidus et la structure 3D de la protéine. En règle générale, elles se font actuellement en utilisant des échantillons protéiques enrichis isotopiquement en ¹⁵N et en ¹³C combinés à des techniques multidimensionnelles. La détermination des résidus qui sont affectés par l'association avec le sucre, une fois reportée sur une structure tridimensionnelle de la protéine permet de définir la zone d'interaction – ce qui est appelé cartographie de déplacement chimique. L'analyse RMN couplée à d'autres techniques permet l'étude des différents modes d'interaction des HS avec leur partenaire (Ziarek et al. 2013) (Poluri et al. 2013) (Saesen et al. 2013). Ces études ont été rendues possibles par la connaissance fine des GAG en RMN à travers de nombreux travaux qui se sont intéressés notamment aux déplacements chimiques des différents résidus issus de différents GAG avec différents motifs de sulfatation, différentes longueurs, différents taux d'épimérisation (Guerrini et al. 2005) (Huckerby et al. 2005) (Pomin 2013)... Si la localisation de la zone d'interaction du côté du partenaire protéique est relativement aisée (dès le moment où on dispose d'un échantillon marqué, soluble, concentré et stable dans le temps), la même étude est extrêmement difficile côté sucre. Cela est en particulier dû au fait que les GAG ou les fragments utilisés dans les études sont d'origine animale et donc les isotopes sont à l'abondance naturelle. De ce fait, il est très difficile de disposer d'informations directes sur les variations induites sur le sucre par la complexation. A l'heure actuelle, seule l'étude de l'interaction de la chimiokine CXCL12 fait appel à des fragments de GAG marqués ¹³C issus d'héparosan modifiés (Laguri et al. 2011) où la position du sucre dans la structure du

complexe est obtenue par des méthodes d'amarrage moléculaire comme la méthode HADDOCK (*High Ambiguity Driven biomolecular DOCKing*) (Figure 64).

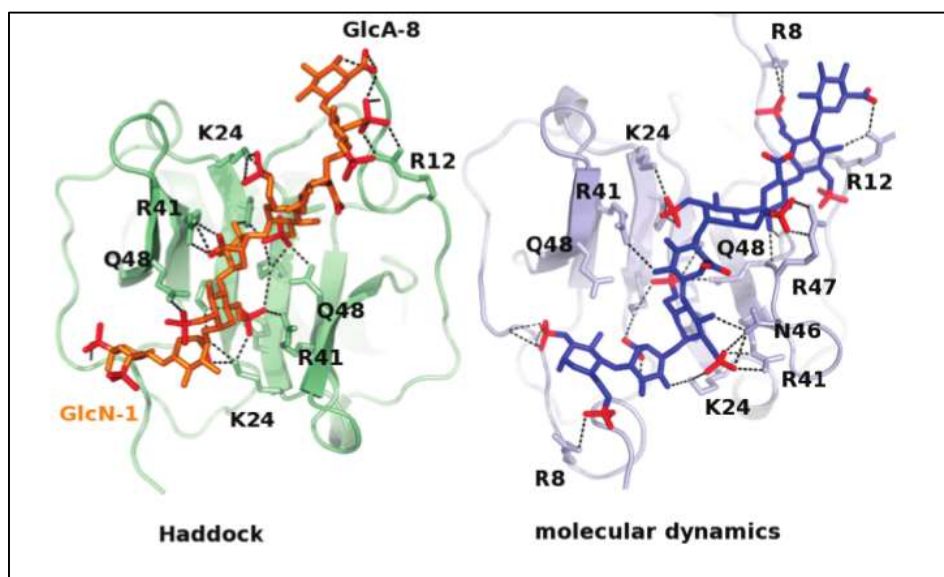


Figure 64 : Comparaison entre deux structures d'interaction CXCL12-dp8 d'héparosan modifié GlcNS,6S obtenu par la méthode d' amarrage moléculaire HADDOCK (gauche) et par dynamique moléculaire (droite). D'après (Laguri et al. 2011).

e) Etudes des GAG par RMN en phase solide

La technique de RMN à l'état solide (ssNMR) a, jusqu'à maintenant, été employée essentiellement en physique des matériaux et commence à se développer dans le domaine de la biologie cellulaire (Renault et al. 2010) (Renault et al. 2012). Les méthodes de RMN en phase solide ont été développées pour l'étude de molécules ou de complexes insolubles, ces méthodes étant indépendantes de la taille de la molécule contrairement aux techniques de RMN dites en solution. Dans ces conditions « solides », les interactions dipolaires ne sont plus moyennées à zéro comme en phase liquide et normalement devraient conduire à une multiplicité et un élargissement important des signaux. Toutefois, les techniques de rotation à l'angle magique (HR-MAS pour *high resolution magic angle spinning* en particulier) où l'échantillon est placé en rotation à un angle de 54,74° par rapport au champ permettent de s'affranchir de ces effets et d'obtenir des spectres où les déplacements chimiques et les largeurs de raies des signaux vont être similaires à ceux qui pourraient être obtenus par RMN en solution.

Les études des GAG par ces méthodes sont en fait très peu nombreuses. On ne compte actuellement qu'une vingtaine d'études faisant appel à cette technique et la plupart sont consacrés à la caractérisation des GAG dans les tissus osseux ou cartilagineux. Dans ces derniers, les GAG peuvent représenter jusqu'à 10 % du poids humide et jouent un rôle essentiel au niveau des forces de compression qui y sont appliquées. Leur caractérisation est une étape importante dans la génération de cartilages synthétiques les plus proches des cartilages naturels. Leur utilisation constituerait une solution à plus long terme pour les personnes souffrant d'arthrose. Par exemple, Scheidt et ses collaborateurs se sont intéressés

à l'analyse comparée de cartilages naturels et issus de la bio-ingénierie. Différents temps de relaxation ont été observés dans les chondroïtines sulfates des deux composés montrant une dynamique différente entre cartilage animal et de laboratoire (Scheidt et al. 2010). Concernant les HS et Hp, le nombre d'étude est encore plus restreint bien que ces GAG puissent interagir avec de nombreuses molécules et former des réseaux insolubles comme c'est le cas au niveau des matrices extra-cellulaires. Ils sont également capables d'interagir avec des protéines fibrillaires comme l'amyloïde-Bêta ($A\beta$) (Madine et al. 2009), la protéine Tau (Goedert et al. 1996), l' α -synucléine (Cohlberg et al. 2002), les prions (Cortijo-Arellano et al. 2008). Certaines formes de ces protéines forment des fibres et des plaques insolubles qui s'accumulent au niveau du cerveau et sont responsables de maladies neurodégénératives. La connaissance du rôle des GAG dans la formation de ces fibrilles est donc très importante sur le plan de la santé humaine. Une première étude par ssNMR de ce type d'interaction a ainsi pu montrer dans le cas de l' $A\beta(1-40)$, qu'une forme dé-*N*-sulfatée/ré-*N*-acétylée d'héparine pouvait interagir avec les fibrilles et même favoriser leur association (Madine et al. 2009). Le site de fixation de l'héparine a été ensuite caractérisé toujours par ssNMR par la même équipe (Madine et al. 2012). Les HS et Hp induisent également la formation de neurofibrilles de la protéine Tau dans le cerveau (Goedert et al. 1996) et constituent également un élément pathologique de la maladie d'Alzheimer. La présence d'HS au sein du cytoplasme constituerait une première étape de la maladie, induisant la formation de fibrilles de tau (Sibille et al. 2006) et réduisant la mobilité de la protéine (Sillen et al. 2005a) (Sillen et al. 2005b). Toutefois, aucune étude par RMN à l'état solide de l'association GAG-tau n'a été publiée jusqu'à maintenant.

Problématique et objectif du travail

Comme nous avons pu le voir, les GAG sont des molécules ubiquitaires qui jouent des rôles extrêmement divers et essentiels dans les différents tissus de l'organisme, au cours du développement et tout au long de la vie. En effet, du fait de leur localisation au niveau des surfaces cellulaires et dans les matrices extracellulaires, ils occupent une place stratégique dans la communication des cellules avec son environnement. Les héparanes sulfates, en particulier, sont capables d'interagir avec un très grand nombre de protéines (plus de 400 protéines recensées à ce jour), et de jouer un rôle actif dans différents processus biologiques. En effet, les HS, de par l'interaction avec leurs ligands protéiques, vont intervenir dans la croissance, la prolifération, la différenciation ou encore la migration cellulaire. De plus, dans un contexte pathologique, ils sont également impliqués dans de nombreuses maladies. Enfin, les HS sont également utilisés par certains pathogènes (virus par exemple) comme récepteurs d'attachement et d'entrée dans la cellule.

La capacité d'interaction des HS prend sa source au niveau de sa diversité structurale, de l'importante combinatoire des modifications des unités disaccharidiques qui les composent et de leur dynamique. Bien que leur structure ne soit pas génétiquement codée, la composition saccharidique des HS est très liée au tissu dans lequel ils sont exprimés et à son état physiologique.

Du fait de cette hétérogénéité structurale, qui caractérise les HS, il n'existe pas de méthode analytique simple et efficace pour les analyser. Les études structurales se sont concentrées principalement sur des fragments oligosaccharidiques très modifiés, issus d'HS entier, permettant de déterminer les modifications critiques pour les interactions HS-protéines. Les données existantes sur les HS cellulaires natifs sont encore très parcellaires. La technique de RMN, malgré sa faible sensibilité, est une technique de choix pour l'étude des polysaccharides car elle permet l'étude de molécules flexibles et d'échantillons complexes selon la résolution.

Afin de permettre l'étude des HS au plus proche des conditions physiologiques cellulaires, je me suis intéressé au cours de ma thèse à l'élaboration d'un protocole efficace de marquage des chaînes d'HS cellulaires à l'isotope ^{13}C du carbone et de purification dans le but de déterminer les caractéristiques structurales de ces composés dans leur état natif par la technique de RMN.

Différentes approches ont été testées, tout d'abord l'étude par RMN liquide des polysaccharides d'HS en solution extraits des membranes cellulaires à partir d'un modèle de cellules épithéliales. Par la suite, en utilisant différents modèles cellulaires, nous avons voulu montré que cette approche pouvait mettre en évidence les modifications structurales des HS au cours de différents événements cellulaires. Je me suis donc intéressé dans un premier temps à l'impact que peut avoir l'absence d'une enzyme de biosynthèse des HS, vitale pour l'organisme, sur la structure des chaînes d'HS, et dans un second temps, à l'évolution de ces structures au cours d'un processus biologique tel que la différenciation cellulaire.

Un autre objectif de ma thèse a été d'étudier ces chaînes de GAG dans leur condition protéoglycanique, au niveau des surfaces cellulaires ou de membranes cellulaires isolées par des techniques de RMN à l'état solide.

Matériel et Méthodes

A. Culture cellulaire

1. Cellules HeLa adhérentes

Les cellules épithéliales HeLa (ECACC) issues de cancer de col de l'utérus sont cultivées dans des conditions adhérentes dans une flasque de 75cm² (Nunc) à 37°C dans une atmosphère humide à 5% de CO₂. Le milieu de culture est composé de DMEM (*Dulbecco's Modified Eagle Medium*) contenant 2mM de glutamine et 4,5g/L de D-glucose (Life Technologies). Il est supplémenté par 10% de sérum de veau fœtal (SVF) (Life Technologies) préalablement inactivé à la chaleur. Les cellules sont maintenues par traitement à la trypsine dès 80% de confluence tous les 4 jours en moyenne.

2. Cellules HeLa en solution et taux de mortalité

15 millions de cellules HeLa sont décollées de la boîte de culture par 3mL de trypsine pendant 10mn à 37°C et sont placées dans une flasque *spinner* à agitation magnétique contenant 20mL de milieu DMEM avec 2mM de Glutamax et 4,5g/L de D-glucose et supplémenté par 10% de SVF inactivé à la chaleur. La rotation appliquée au sein du *spinner* est ménagée afin de prévenir la sédimentation des cellules ainsi que leur destruction. Les cellules sont régulièrement comptées afin de déterminer le taux de mortalité cellulaire dans ces conditions.

Pour le comptage et l'estimation du taux de mortalité, un aliquot de cellules est mis en présence de bleu de trypan pour marquer les cellules apoptotiques. Le nombre de cellules mortes (marquées au bleu) et vivantes (non marquées) sont ensuite comptées sur cellule de Malassez dans un volume de 0,01μL. Ce nombre est rapporté en nombre de cellules par millilitre de culture. Le taux de mortalité est calculé selon la formule :

$$\frac{\text{Nombre de cellules marquées au bleu de trypan}}{\text{Nombre de cellules totales comptées}} \times 100$$

3. Fibroblastes de souris embryonnaires (MEF)

Les fibroblastes d'embryon de souris ou MEF sauvages ou *KO* au gène *ndst-1* (protocole Ringvall et al. 2000) (don de nos collaborateurs du laboratoire de Léna Kjellen à l'université d'Uppsala, Suède) sont cultivées à 37°C dans des conditions adhérentes dans une flasque de 75cm² à 37°C dans une atmosphère humide à 5% de CO₂. Le milieu de culture est composé de DMEM contenant 2mM de Glutamax et 4,5g/L de D-glucose. Il est supplémenté par 10% de sérum de veau fœtal préalablement inactivé à la chaleur ainsi que d'un cocktail de 100U/mL de pénicilline et de 100μg/mL de streptomycine (Life Technologies). Les cellules sont maintenues par traitement à la trypsine dès 80% de confluence tous les 4 jours en moyenne.

4. Cellules Caco-2

Les cellules Caco-2 (ECACC) issues d'adénocarcinome de colon sont cultivées dans des conditions adhérentes dans une flasque de 75cm² à 37°C dans une atmosphère humide à 5% de CO₂. Le milieu de culture est composé de DMEM contenant 2mM de Glutamax et 4,5g/L de D-glucose (Life Technologies). Il est supplémenté par 10% de sérum de veau fœtal (Life Technologies) préalablement inactivé à la chaleur ainsi que de 1% d'acides aminés non-essentiels (NEAA) (Life Technologies) et d'un cocktail de 100U/mL de pénicilline et de 100µg/mL de streptomycine (Life Technologies). Les cellules sont passées à 80% de confluence tous les 4 jours en moyenne.

La différenciation des cellules Caco-2 s'opère spontanément lorsque les cellules arrivées à confluence sont maintenues ainsi durant une période de 12 jours de culture dans ces conditions, le milieu de culture étant renouvelé tous les 4 jours (Sadir et al. 2004).

5. Marquage au ¹³C et amplification des cellules

a) Marquage

Le marquage métabolique des cellules HeLa, Caco-2 et MEF au ¹³C-glucose s'effectue lors de l'amplification des cellules sur une surface cellulaire d'environ 1800cm². Les différents types cellulaires sont cultivés de 10 % à 100 % confluence dans les mêmes conditions et dans des milieux de composition similaire à ceux précédemment décrits dépourvus en D-glucose et pyruvate (PAA Laboratories ou Life Technologies) et supplémentés en ¹³C-glucose (Cambridge Isotope Laboratories) à 4,5g/L. Les cellules sont récupérées à confluence pour analyse.

Pour les cellules Caco-2, elles sont maintenues à confluence pendant 12 jours afin que s'effectue complètement le processus de différenciation dans le milieu marqué renouvelé deux fois.

b) Flasques de culture

-T600 (EMD Millipore) : 10 % des cellules sont expandues dans trois flasques T600 (surface de culture = 1800cm²) avec 60mL (V_{total}=180mL) de milieu marqué dans chacun des flasques équitablement répartis sur les différents niveaux (Figure 65A). Les cellules confluentes sont ensuite récoltées.

-HYPERflaskTM (Corning[®]) : 10 % des cellules sont expandues dans une HYPERflaskTM (surface de culture = 1750cm²) avec 500mL de milieu marqué (Figure 65B). Les cellules confluentes sont ensuite récoltées.

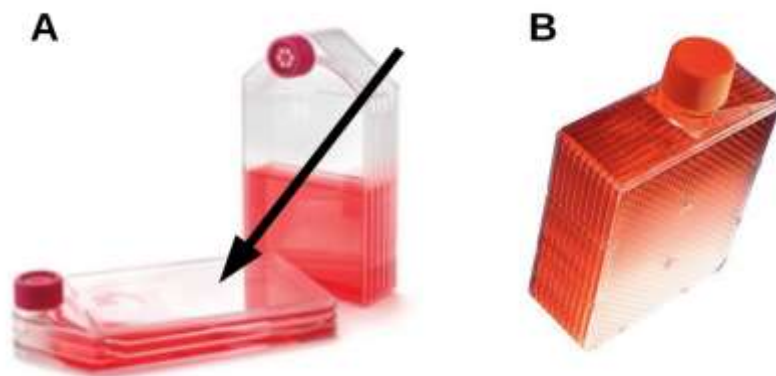


Figure 65 : Différents systèmes de culture cellulaire adaptés pour une production importante à partir de cellules adhérentes. Système T600 de Millipore (A, flèche) et HYPERflask™ de Corning® (B).

B. Cytométrie en flux

1. Préparation des cellules

L'expression en GAG à la surface des cellules est analysée par cytométrie en flux. Dans un premier temps, les cellules HeLa et MEF sont décollées de leur support à l'aide d'une solution d'EDTA (versène, Life Technologies), qui va chélater les ions métalliques et notamment l'ion calcium (Ca^{2+}) nécessaire pour le rôle d'adhésion des « molécules d'adhésion cellulaire » (CAM) : glycoprotéines (intégrines, cadhérines)... Les cellules Caco-2 quant à elles sont traitées à l'accutase (Life Technologies) 10mn à 37°C afin de détacher les cellules de leur support et de les dissocier entre elles. La digestion est stoppée par ajout de milieu de culture.

Une fois détachées, les cellules sont lavées à 4°C dans du PBS et reprises dans du PBS additionné de 1mM EDTA, 1% BSA, 0,002% azide.

Les cellules sont incubées en présence d'un anticorps IgM primaire monoclonal « 10E4 » (10µg/mL ; Seikagaku Corporation) dirigé contre un épitope comprenant une GlcNS spécifique des HS ou d'un anticorps IgM primaire monoclonal « CS-56 » spécifique des chaînes de CS-A et CS-C (dilué au 1/100^e ; Sigma Aldrich) ou d'un anticorps isotype IgM (10µg/mL ; Chemicon, Millipore) pendant 1h à 4°C. Les cellules sont ensuite lavées 3 fois dans le tampon puis incubées avec un anticorps secondaire (1/300^e, Interchim) reconnaissant l'anticorps primaire. L'anticorps secondaire est couplé à une molécule fluorescente FITC (Fluorescein IsoThioCyanate) qui va être détectée par cytométrie en flux. Les cellules sont lavées 3 fois avec du tampon et fixées pendant 10mn dans 4% paraformaldéhyde (PFA) afin de figer les cellules et leur surface cellulaire pour leur analyse.

2. Analyse

Les cellules passant dans le cytomètre en flux vont être isolées une à une avant de passer devant un faisceau laser dont la lumière va être déviée en fonction de la taille de la cellule et de sa complexité (granularité...) cytoplasmique. En effet, un détecteur en face de la source laser va capter l'information de diffraction de la lumière axiale (FSC pour *Forward-*

Scattered light) tandis que des capteurs situés à 90° de la source laser vont enregistrer la réfraction de la lumière sur les particules internes des cellules (*SSC pour Side-Scattered light*) (Figure 66). Ces deux informations sont traitées et différentes populations peuvent être sélectionnées en fonction de leur taille et de leur granularité. Les cellules en apoptose (qui se caractérisent par leur granularité importante du fait de la condensation et fragmentation de leur ADN, donc avec un SSC important) et les corps apoptotiques (ayant une taille bien inférieure à une cellule saine, donc avec un FSC faible) ne sont pas pris en compte pour l'analyse. Une fois la population de cellules sélectionnée, le fluorochrome est excité à sa longueur d'onde d'absorption (ici 488nm pour le FITC) et sa fluorescence est lue à sa longueur d'onde d'émission (530nm pour le FITC). Le résultat obtenu en sortie est la quantité de fluorescence par cellule. Cette fluorescence est proportionnelle à la quantité de fluorochrome fixé sur la cellule.

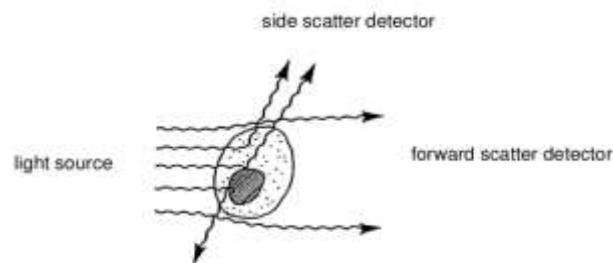


Figure 66 : Schéma de la diffraction et réfraction de lumière par une cellule

C. Digestions enzymatiques...

1. ...des surfaces cellulaires aux protéases

a) Trypsine

Les surfaces de 200 millions de cellules cultivées en milieu marqué, et préalablement rincées avec 60mL de PBS, sont digérées dans 40mL de solution de trypsine (Life Technologies) pendant 40mn à 37°C à pH neutre. La trypsine est une endoprotéase qui va couper les protéines du côté C-terminal des acides aminés basiques lysine et arginine à condition que l'acide aminé suivant en C-ter ne soit pas une proline. Le digestat cellulaire est récupéré dans 150mL de PBS et centrifugé à 1000g pendant 10mn à température ambiante. Le culot contenant les cellules digérées est séparé du surnageant contenant les GAG en solution. Le surnageant est ensuite centrifugé une seconde fois à 3500g pendant 10mn à température ambiante. Seul le surnageant est conservé pour la prochaine étape de purification par chromatographie.

b) Papaïne

Les surfaces de 200 millions de cellules cultivées en milieu marqué, et préalablement rincées avec 60mL de PBS, sont digérées avec 1mg/mL de papaïne (Sigma Aldrich) à 65°C

pendant 16h dans un tampon composé de 100mM acétate de sodium, 5mM EDTA, 5mM cystéine, pH 6,8 dans un volume de 24mL. La papaïne est une endoprotéase qui va couper préférentiellement les protéines du côté C-terminal au niveau des acides aminés basiques lysine et arginine à condition que l'acide aminé adjacent en position C-terminale ne soit pas une valine et que l'acide aminé en position N-terminale de la lysine/arginine soit hydrophobe. Le digestat est récupéré dans 120mL de PBS et centrifugé à 1000g pendant 10mn à température ambiante. Le culot contenant les cellules digérées est séparé du surnageant contenant les GAG en solution. Le surnageant est ensuite centrifugé une seconde fois à 3500g pendant 10mn à température ambiante. Seul le surnageant est conservé pour la prochaine étape de purification par chromatographie.

c) Pronase

200 millions de cellules cultivées en milieu marqué, et préalablement rincées dans 60mL de PBS, sont d'abord décollées de leur support avec 36mL de versène (Life Technologies) pendant 30mn à 37°C. Les cellules récupérées en solution sont centrifugées à 1000g pendant 5mn à température ambiante, et le culot cellulaire est repris dans 20mL de tampon 100mM Tris, 1mM CaCl₂, pH 8 à 37°C auquel est ajouté 1mg/mL de cocktail d'enzymes « pronase » (Roche). La pronase contient plusieurs endoprotéases et exoprotéases peu spécifiques permettant de digérer généralement les protéines en leur unité d'acides aminés les plus simples. Après 40mn la solution est centrifugée (1000g, 5mn, température ambiante) et le culot cellulaire est digéré une seconde fois dans 20mL de tampon d'enzyme avec l'enzyme pendant 40mn. La solution est centrifugée (1000g, 5mn, température ambiante). Les surnageants sont rassemblés et centrifugés une nouvelle fois à 3500g pendant 10mn à température ambiante et seul le surnageant est conservé pour la prochaine étape de purification par chromatographie.

2. ...des extraits cellulaires aux nucléases

a) Benzonase

Les échantillons lyophilisés sont repris dans 100µL de tampon benzonase (5mM MgCl₂, pH 9) contenant l'enzyme (Millipore) à 2500U/mL pendant 21h à 37°C. La benzonase est une endonucléase qui dégrade en 3' toutes les formes d'ADN et d'ARN en oligonucléotides de 2 à 5 bases sous leur forme 5' phosphate. Les échantillons ainsi digérés sont ensuite prêts pour la prochaine étape de purification par chromatographie.

b) DNase I

Les échantillons lyophilisés sont repris dans 20000U de DNase I (Roche) dans du tampon 150mM NaCl, 100mM Tris, 10mM MgCl₂, pH 8 à 37°C pendant 24h avec ajout d'enzyme et en tampon frais au bout de 16h. La DNase I est une endonucléase qui clive spécifiquement les ADN au niveau des liaisons phosphodiester des cytidines et thymidines

laissant des extrémités nucléotidiques 3'OH. En fin de digestion, l'enzyme est désactivée par ajout de 25mM EDTA. Les échantillons ainsi digérés sont ensuite prêts pour la prochaine étape de purification par chromatographie.

D. Purification par chromatographie échangeuse d'anions

1. Chargement des échantillons

a) Après digestion aux protéases

Les surnageants de centrifugation correspondant aux molécules solubles digérées des surfaces cellulaires sont chargées sur une colonne de chromatographie contenant 4mL de matrice cationique Diéthylaminoéthyl-Sephacel (DEAE-Sephacel) (Sigma-Aldrich) à un débit de 0,5mL.mn⁻¹. La colonne est ensuite connectée à un système de purification FPLC (*Fast Protein Liquid Chromatography*) AKTÄ pour leur élution.

b) Après digestion aux nucléases

Les échantillons sont repris dans du tampon 20mM Na₂HPO₄/NaH₂PO₄, 150mM NaCl, pH 6,8 (qsp 2mL) et sont injectés sur une colonne de chromatographie contenant 4mL de matrice DEAE-Sephacel (Sigma-Aldrich) reliée à un système de purification FPLC AKTÄ pour leur élution.

Préalablement à chaque charge/injection sur la matrice DEAE, celle-ci est pré-équilibrée dans 5 volumes de colonne avec un tampon 20mM Na₂HPO₄/NaH₂PO₄, 150mM NaCl, pH 6,8 (tampon A). Après charge/injection de l'échantillon, 10mL de ce tampon A sont passés au travers de la colonne de chromatographie afin d'éluer les molécules n'interagissant pas avec la matrice (molécules neutres, cationiques) ou très faiblement, la fraction constituant l'effluent ou *flowthrough*.

L'élution des molécules (poly)anioniques interagissant plus fortement avec la résine cationique est réalisée selon deux procédés.

2. Elution en *step* de sel

Cette méthode d'élution consiste en un changement de concentration en NaCl soudain d'un tampon A (20mM Na₂HPO₄/NaH₂PO₄, 150mM NaCl) en un tampon B 20mM Na₂HPO₄/NaH₂PO₄, 1,5M NaCl à 0,5mL.mn⁻¹. Toutes les interactions électrostatiques sont alors cassées et toutes les molécules d'interaction sont éluées dans ce tampon B. L'élution des molécules interagissant avec la matrice est suivie au travers de l'absorbance des fractions à 260nm pour les acides nucléiques et à 215nm et 280nm pour les protéines. La concentration en NaCl est suivie à travers la conductivité de l'éluat (en milliSiemens par centimètre mS/cm).

3. Elution en gradient de sel

Cette méthode d'élution consiste en un changement progressif et linéaire du tampon A vers le tampon B, et donc d'une augmentation de la concentration en sel. Ce gradient 0-100% en tampon B se déroule sur 20mL et permet d'éluer différentes fractions selon leur force d'interaction électrostatique avec la matrice. Les molécules de plus faible interaction (faible charge négative) seront éluées en une concentration de sel plus faible que les molécules de forte interaction (charge nette très négative), nécessitant une concentration en NaCl plus importante pour casser toutes les interactions électrostatiques. La concentration en sel est suivie à travers la conductivité de l'éluat et est exprimée en mS/cm. L'élution des contaminants nucléiques et protéiques est suivie à travers leur absorbance à la longueur d'onde de 260nm et 280nm respectivement ainsi qu'à 215nm (absorbance des liaisons peptidiques).

Les fractions d'élution sont ensuite dessalées pour la suite des analyses.

E. Préparation des extraits d'HS

1. Dessalage

Les fractions obtenues à l'issue des chromatographies échangeuses d'anions et des digestions cellulaires présentent des concentrations en sel variées. Pour la suite de la purification ou pour leurs analyses par RMN, l'absence de NaCl est requise. Différentes techniques ont été utilisées.

a) Système PD-10

2,5mL d'échantillon salé d'intérêt sont élués dans de l'eau pure par gravité sur colonne de chromatographie d'exclusion de taille (PD-10, GE Healthcare) contenant une résine de type Sephadex G-25. Les molécules de masse molaire relative (M_r) supérieure à 5000 sont séparées des plus petites molécules comme le NaCl. Ainsi, 3,5mL d'éluat dessalé sont récupérés pour chaque échantillon de 2,5mL.

b) Système Hitrap Sephadex

Le système Hitrap (GE Healthcare) contient la même matrice que celle des colonnes PD-10 précédemment citée et repose sur les différents temps d'élution entre macromolécules et petites molécules. Le système Hitrap est relié à un appareil FPLC et permet de déterminer un débit d'élution de 2,5mL.mn⁻¹. 2mL d'échantillon sont injectés dans le système par dessalage et 5mL d'éluat d'intérêt est récupéré à la fin de chaque procédé. L'élution des macromolécules peut être suivie à différentes longueurs d'onde (280nm, 260nm, 215nm) et comparée avec le temps d'élution du NaCl suivi par la conductivité de l'échantillon.

c) Dialyse

Les échantillons issus des digestions protéases et nucléases lyophilisés et repris dans 30mL d'eau sont placés au sein d'une membrane de dialyse poreuse fermée laissant diffuser les molécules d'une taille inférieure à 1000Da (SpectrumLabs) dans 5L d'eau ultra-pure. Les bains de dialyse sont renouvelés 5 fois. Ce système de dessalage est utilisé pour la mise en place d'un protocole visant à réduire le nombre d'étapes par chromatographie échangeuse d'anions (voir Résultats).

Les échantillons dessalés par les différentes méthodes utilisées ont un volume plus important que le volume d'extrait salé. Afin de pouvoir se placer dans des conditions de purification ou d'analyse optimales, les échantillons obtenus après dessalage sont lyophilisés.

2. Lyophilisation

Les échantillons sont préalablement congelés puis placés dans un lyophilisateur (Alpha 1-2 LDplus, Bioblock Scientific) sous vide à -50°C afin de sublimer l'eau (passage de l'état solide à l'état gazeux) qui sera ensuite piégée. L'échantillon est ainsi déshydraté.

3. Préparation des HS avec réduction d'étapes pré-analytiques

Lors de cette préparation, les extraits solubles des digestions membranaires à la trypsine sont entièrement lyophilisés. L'extrait sec est repris dans 30mL d'eau. Les éléments précipités sont éliminés par centrifugation (3500g, 10mn, température ambiante) et le surnageant est dialysé comme précédemment décrit. Le dialysat est ensuite lyophilisé.

F. Dosage des HS

Après extraction des HS des surfaces cellulaires et purification par chromatographie échangeuse d'anions, différentes techniques de dosage des HS sont testées.

1. Dosage par carbazole

La méthode de dosage au carbazole (Bitter & Muir 1962) repose sur la réaction des acides uroniques transformés en acides formylfuroïques à chaud par l'acide sulfurique qui vont se condenser avec le colorant carbazole absorbant à une longueur d'onde spécifique de 530nm. La détermination de la concentration en acides uroniques est réalisée par rapport à une gamme étalon d'HS dont la concentration est connue au préalable.

250µL de 0,25M de tétraborate de sodium (dans acide sulfurique H₂SO₄) sont placés dans un tube en verre froid. 50µL d'échantillon sont ajoutés précautionneusement à la surface, puis le tube fermé est agité d'abord lentement puis rapidement à 4°C (la réaction est fortement exothermique). Le tube est ensuite incubé 10mn à 100°C puis refroidi à température ambiante. 10µL de carbazole sont ajoutés au volume réactionnel, puis le tube est mis sous agitation et incubé 15mn à 100°C avant d'être à nouveau refroidi à température

ambiante. La densité optique (DO) est lue à 530nm à laquelle est soustraite son absorbance à 920nm (qui correspond à une interférence causée par une oxydation).

La DO de l'échantillon est reportée sur la droite obtenue par les valeurs de la gamme étalon de différentes concentrations en HS dilués en série (0-6,25-12,5-25-50 $\mu\text{g.mL}^{-1}$).

2. Dosage par Blyscan™

Les échantillons sont repris dans 100 μL d'eau auquel est ajouté le réactif Blyscan (Figure 67) puis agité mécaniquement pendant 30mn. Au cours de ces 30mn, le réactif va interagir avec les GAG sulfatés (sGAG) et précipiter avec ceux-ci. Les échantillons sont ensuite centrifugés à 13400g pendant 10mn. Le surnageant est éliminé et le culot est séché. 500 μL de réactif de dissociation est ensuite ajouté au culot. Le mélange est ensuite vortexé pour libérer le colorant en solution. L'absorbance de la solution est ensuite lue à 630nm. Les résultats sont reportés sur une droite obtenue avec les absorbances d'une gamme étalon en sGAG de 0-1-2-3-4-5 μg pour déterminer la quantité en sGAG dans l'échantillon. Les sGAG fournis dans le kit sont des C4S.

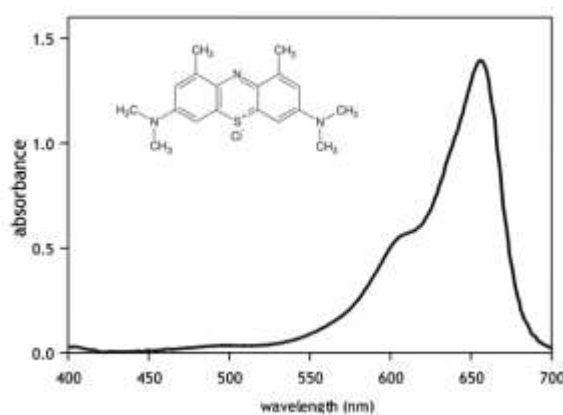


Figure 67 : Structure et spectre d'absorbance du colorant Blyscan™

3. Dosage par RMN

La détermination de la quantité d'HS d'un échantillon inconnu est réalisée à l'aide d'un échantillon de masse connue.

a) Analyse de l'échantillon de référence

2mg d'HS commercial (Seikagaku) issus de rein de bœuf sont analysés par RMN dans une expérience de corrélation ^1H - ^{13}C HSQC à 37°C pendant 13h15mn dans 200 μL de D_2O (99,9 % D_2O , Sigma Aldrich) dans un tube RMN de 3mm de diamètre. La résolution est de 11 Hz en dimension directe (^1H), de 176 Hz en dimension indirect (^{13}C) et chaque expérience est répétée 180 fois pour améliorer le rapport signal sur bruit. Du fait de l'abondance naturelle en ^{13}C , seul 1,1% des noyaux sont observables, ce qui équivaut à 20 μg marqué pour un échantillon contenant au total 2mg.

Les signaux des pics de corrélation H₂-C₂ des glucosamines étant bien résolus, l'intégration du volume de ces pics est représentatif de la quantité en HS.

Pour un temps expérimental de 13h15 : $V_{\text{Glucosamines}} \propto \text{masse} \times t_{\text{exp}}$ où $V_{\text{Glucosamines}}$ est le volume des pics de corrélation H₂-C₂ des Glucosamines et t_{exp} le temps expérimental.

b) Analyse de l'extrait d'HS

L'échantillon d'HS de masse inconnue est analysé également par RMN pendant 15h56mn (résolution directe de 11 Hz, résolution indirecte de 26Hz pour une fenêtre spectrale de 21kHz et expérience accumulée 32 fois). Les volumes des pics sont intégrés pour les corrélations H₂-C₂ des glucosamines.

La masse de l'échantillon est déterminée selon l'équation : $m_{\text{éch}} = m_{\text{ref}} \times \frac{V_{\text{éch}}/t_{\text{éch}}}{V_{\text{ref}}/t_{\text{ref}}}$ où

m_{ref} est la masse de l'échantillon de référence observable (soit 20μg), V_{ref} est la somme des volumes des corrélations H₂-C₂ des glucosamines de l'échantillon de référence, t_{ref} est le temps d'analyse de l'échantillon de référence, $m_{\text{éch}}$ est la masse d'HS de l'extrait cellulaire à déterminer, $V_{\text{éch}}$ est la somme des volumes des corrélations H₂-C₂ des glucosamines de l'échantillon issu des cellules, $t_{\text{éch}}$ est le temps d'analyse de l'extrait cellulaire.

G. Analyse par RMN liquide des extraits d'HS

Tous les échantillons analysés par RMN sont échangés plusieurs fois dans 99 % D₂O avant analyse pour éliminer les échanges des protons résiduels de l'eau. Les échantillons sont solubilisés dans 200μL de D₂O 99 % et placés dans un tube RMN de 3mm de diamètre.

1. Expériences de corrélation

Tous les extraits d'HS issus des digestions des membranes sont analysés par des techniques RMN de corrélation ¹H-¹³C HSQC (*Heteronuclear Single Quantum Coherence*) (d'après (Bodenhausen & Ruben 1980)) : la séquence permet de déterminer les déplacements chimiques des couples de noyaux ¹H-¹³C vicinaux. Chaque couple va alors donner un pic de corrélation dépendant de la fréquence de résonance du proton et du carbone associé. Les temps d'expériences sont généralement compris entre 13h et 16h avec un nombre d'itération de 180 en moyenne et d'une résolution dans la dimension indirecte (¹³C) de 100-200Hz. Les périodes de transfert ¹H-¹³C sont calculées à partir d'un couplage scalaire J_{CH} de 145Hz. Les échantillons sont analysés à 37°C dans des spectromètres 600MHz ou 700MHz équipés de cryosondes.

Les déplacements chimiques ont été référencés par rapport au signal de l'eau. Les déplacements chimiques ¹³C ont été référencés indirectement en utilisant le rapport des rapports gyromagnétiques entre proton et hétéroatome (¹³C/¹H=0.251449530) comme décrit par Wishart (Wishart et al. 1995).

2. Quantification de la modification des HS

La quantification des différents taux de modification est réalisée grâce à l'intégration des volumes des pics résolus permettant de déterminer : la *N*-sulfatation, l'épimérisation, la 2-*O*-sulfatation, la 6-*O*-sulfatation ainsi que d'estimer la longueur des chaînes.

L'intégration des pics est réalisée à l'aide d'un programme (langage python) mis en place et disponible au laboratoire.

H. Etudes des cellules entières

1. Préparation des cellules

a) Cellules entières

Pour toutes les études par RMN du solide, les cellules sont décollées de leur support de culture par du versène (30mL pour une surface de culture de 1800cm²). Les cellules sont ensuite manipulées à 4°C, centrifugées (1000g, 5mn) et rincées deux fois dans 15mL de PBS. Les cellules sont ensuite placées dans un rotor 3,2mm (25 millions) ou 7mm (200 millions) pour leurs analyses par RMN solide par courtes centrifugations (2000g).

b) Cellules dépourvues d'HS, CS et N-glycans

Les cellules récupérées et culottées comme décrit précédemment sont repris dans 250µL d'une solution enzymatique RPMI (Roswell Park Memorial Institute) 1640, 5 % SVF, 2mM CaCl₂ contenant 0,1U/mL d'héparinase I, 2U/mL d'héparinase III (Grampian Enzymes) (pour spécificité des coupures voir Figure 39), et 500mU de chondroïtinase ABC (Sigma Aldrich). Les cellules sont incubées à 37°C pendant 90mn sous agitation dans cette solution. Une partie de ces cellules sont récupérées par centrifugation, lavées en PBS et placées dans un rotor pour analyse comme précédemment.

Une seconde partie des cellules est reprise dans 200µL de tampon 200mM Na₂HPO₄/NaH₂PO₄ pH 8 contenant 10U/mL de *N*-glycosidase F (Roche) coupant les *N*-glycosides pendant 2h à 37°C. Après lavage en PBS les cellules sont placées dans un rotor pour analyse par RMN solide.

2. Analyse par RMN solide

a) Par rotation à l'angle magic MAS (Magic Angle Spinning)

Différentes expériences sont enregistrées :

-1D par excitation directe du ¹³C : Les échantillons de cellules sont placés dans un rotor 3,2mm à l'angle magique (54,74°) dans un spectromètre 600MHz équipé d'une sonde solide. Les spectres sont obtenus à 12°C et permettent d'observer toutes les résonances carbone détectables.

-Corrélation ^{13}C - ^{13}C via le couplage scalaire par excitation directe du carbone : Ces expériences permettent de corrélérer deux ^{13}C voisins par le couplage scalaire (*via* les liaisons chimiques). Les échantillons de cellules sont placés dans un rotor de 3.2 mm pour la sonde solide du 600 MHz et dans un rotor de 7mm pour la sonde solide du spectromètre 400MHz. Les spectres sont enregistrés à 4°C pendant 15h à une rotation de l'échantillon de 12kHz pour le rotor 3.2 mm et 3,4Khz pour le rotor 7mm. Ces analyses ont été réalisées à l'IBS pour le 600 MHz et à Institut Nanosciences et Cryogénie pour le 400 MHz.

b) Par polarisation dynamique nucléaire (DNP)

Les cellules HeLa sont également analysées par DNP avec des expériences de type DQ (*Double Quanta*) en collaboration avec le Dr Takahashi et le Dr Hediger de l'Institut Nanosciences et Cryogénie. Brièvement, l'aimantation d'un électron non-apparié (appelé radical) d'un agent polarisant, ici le TOTAPOL (1-(TEMPO-4-oxy)-3-(TEMPO-4-amino)propan-2-ol), excité par une émission micro-onde va être transférée vers les noyaux du matériel à analyser. L'aimantation initiale étant beaucoup plus forte, le signal en sortie sera plus important. Ces expériences sont réalisées par le Dr Takahashi sur des échantillons en rotation à l'angle magique placés à 4°C dans un spectromètre 400MHz pendant 15h.

I. Etudes des membranes plasmiques isolées

1. Préparation des membranes

200 millions de cellules cultivées en milieu marqué et préalablement rincées au PBS sont détachées de leur support dans 40mL de versène puis récupérées dans 100mL de PBS. Les cellules sont ensuite centrifugées (1000g, 5mn, 4°C) et placées dans 1,5mL de tampon hypotonique 10mM Tris pendant 45mn à 4°C en présence d'antiprotéases (Roche). Les cellules sont ensuite cassées mécaniquement à l'aide d'un potter (30 coups). 5mM d'EDTA sont ajoutés au lysat cellulaire qui est centrifugé à 4500g, 10mn, 4°C. Le surnageant contenant les membranes plasmiques est gardé et le culot est lavé une deuxième fois dans le tampon Tris et est centrifugé une nouvelle fois dans les mêmes conditions. Les surnageants sont ensuite ultracentrifugés 175000g, 30mn, 4°C pour sédimenter les membranes plasmiques. Le surnageant est une seconde fois ultracentrifugé dans les mêmes conditions. Les membranes plasmiques ainsi sédimentées peuvent être analysées par RMN en phase solide ou être solubilisées dans des détergents afin de réaliser des expériences par RMN en phase liquide.

2. Etudes par ssNMR

Différentes expériences sont réalisées sur les préparations de membranes des cellules HeLa :

-1D par excitation directe du ^{13}C : L'expérience permet d'observer tous les ^{13}C . Les échantillons de cellules sont placés dans un rotor 3,2mm à l'angle magique (54,74°) dans un spectromètre

600MHz équipé d'une sonde solide. Les spectres sont enregistrés à 8°C pendant 11mn sur un échantillon subissant une rotation de 12kHz.

-1D ^{13}C INEPT (*Insensitive Nuclei Enhanced by Polarization Transfer*) : L'expérience permet d'observer les ^{13}C de régions relativement flexibles en transférant l'aimantation du ^1H vers le ^{13}C *via* le couplage scalaire. Les échantillons de cellules sont placés dans un rotor 3,2mm à l'angle magique dans un spectromètre 600MHz équipé d'une sonde solide. Les spectres sont enregistrés à 8°C pendant 3mn sur un échantillon subissant une rotation de 12kHz.

-1D ^{13}C CP (*Cross Polarization*): L'expérience permet d'observer les ^{13}C de régions relativement rigides en transférant l'aimantation du proton au carbone *via* le couplage dipolaire (à travers l'espace). Plus la molécule est rigide, plus le couplage dipolaire est important. Les échantillons de cellules sont placés dans un rotor 3,2mm à l'angle magique dans un spectromètre 600MHz équipé d'une sonde solide. Les spectres sont enregistrés à 8°C pendant 31mn sur un échantillon subissant une rotation de 12kHz.

-Corrélation ^{13}C - ^{13}C via le couplage scalaire par excitation directe du carbone : Ces expériences permettent de corréler deux ^{13}C voisins. Les échantillons de cellules sont placés dans un rotor 3,2mm à l'angle magique dans un spectromètre 600MHz équipé d'une sonde solide. Les spectres sont enregistrés à 8°C pendant 37h à une rotation de l'échantillon de 12kHz.

-2D INEPT : Ces expériences permettent de corréler une paire de noyaux ^1H et ^{13}C directement couplés par une liaison chimique. Les échantillons de cellules sont placés dans un rotor 3,2mm à l'angle magique dans un spectromètre 600MHz équipé d'une sonde solide. Les spectres sont enregistrés à 8°C pendant 28h à une rotation de l'échantillon de 12kHz.

3. Solubilisation des membranes

Les culots membranaires peuvent également être solubilisés dans des détergents pour des analyses par RMN liquide :

a) Triton X-100

Le culot de membranes plasmiques est repris dans 200 μL de détergent non-ionique de 1% Triton X-100 (Euromedex) dans du D_2O à température ambiante et est solubilisé par pipetages successifs à température ambiante.

b) OctylPOE

Le culot de membranes plasmiques est incubé dans 200 μL de solution 3% octylPOE (n-octylpolyoxyéthylène) (D_2O) pendant 1h à 37°C. La solubilisation totale de l'échantillon est vérifiée par l'absence de culot après une ultracentrifugation (175000g, 30mn, 4°C).

c) DPC

Le culot contenant les membranes plasmiques est incubé dans 200 μ L de détergent zwitterionique de DPC/FOS12 (n-dodécylphosphocholine) (D₂O) pendant 2h à 4°C. La solubilisation totale de l'échantillon est vérifiée par l'absence de culot après une ultracentrifugation (175000g, 30mn, 4°C).

4. Digestion des HS et CS

Le culot membranaire peut être repris dans du tampon 50mM NaCl, 5mM Tris, 2mM CaCl₂ (D₂O) pH 7,8 auquel sont ajoutées les enzymes héparinase I, héparinase III (Figure 39) et chondroïtinase ABC à 0,1U.mL⁻¹ finale pour les héparinases I et III et 2U.mL⁻¹ pour la chondroïtinase. La réaction est effectuée à 37°C sous agitation pendant 1h30 et est stoppée par dénaturation thermique des enzymes à 100°C pendant 10mn. L'échantillon est ensuite centrifugé (4000g, 5mn, température ambiante) et le surnageant est placé dans un tube RMN 3mm pour son analyse spectrométrique.

5. Analyse par RMN liquide

a) HSQC

-Analyse des membranes solubilisées non-digérées : Les expériences sont enregistrées sur un spectromètre 600MHz équipé d'une cryosonde pendant 1h15mn (8 accumulations de l'expérience, résolution ¹H de 11Hz et ¹³C de 82Hz pour une fenêtre spectrale ¹³C de 21kHz).

-Analyse des membranes solubilisées digérées aux héparinases I, III et chondroïtinase ABC : Les spectres issus des échantillons digérés sont enregistrés sur un spectromètre 600MHz équipé d'une cryosonde pendant 10h10mn (32 accumulations de l'expérience, résolution ¹H=11Hz ; ¹³C=41Hz pour une fenêtre spectrale ¹³C de 21kHz).

b) HCCH-COSY (COrrrelation Spectroscopy)

Les spectres issus des membranes solubilisées dans l'octylPOE sont analysés par des expériences de type HCCH-COSY. Elles permettent de corrélérer un proton (¹H) avec le carbone (¹³C) sur lequel il est directement lié, et également sur le premier carbone voisin. Les expériences sont enregistrées sur un spectromètre 600MHz équipé d'une cryosonde pendant 39mn (4 accumulations de l'expérience, résolution ¹H de 11Hz et ¹³C de 98Hz pour une fenêtre spectrale de 25kHz).

c) HCCH-TOCSY (TOtal COrrrelation Spectroscopy)

Les spectres issus des membranes solubilisées dans l'octylPOE sont analysés par des expériences de type HCCH-TOCSY. Elles permettent de corrélérer un proton (¹H) avec le carbone (¹³C) sur lequel il est directement lié, ainsi qu'avec tous les carbones voisins séparés entre eux

par une seule liaison chimique. Les expériences sont enregistrées sur un spectromètre 600MHz équipé d'une cryosonde pendant 13h15mn (80 accumulations de l'expérience, résolution ^1H de 11Hz ^{13}C de 98Hz pour une fenêtre spectrale de 25kHz).

Résultats et Discussion

La connaissance des structures des HS au niveau des membranes plasmiques n'étant que parcellaire, le but de notre étude a été de mettre en place un protocole rendant possible l'étude de ces polysaccharides dans un contexte nouveau visant à déterminer leur structure directement à partir de chaînes entières. Dans ces conditions, la RMN est la seule technique actuelle permettant d'analyser les chaînes d'HS en solution. La technique permet également d'envisager l'étude des HS au niveau même des membranes plasmiques.

Nous avons donc, dans un premier temps, vérifié l'expression des HS à la surface de cellules que nous avons choisi comme modèle cellulaire. A partir de celui-ci, des techniques de production d'HS et d'optimisation ont été mises en œuvre. Afin de permettre une analyse par RMN et du fait de la faible abondance naturelle de l'isotope ^{13}C du carbone observable par cette technique, les chaînes d'HS ont été marquées métaboliquement avec le ^{13}C par incorporation de ^{13}C -glucose dans le milieu de culture.

Les HS marqués sont extraits de la surface cellulaire par digestion protéique puis purifiés par chromatographie échangeuse d'anions. Cette étape de purification a nécessité de nombreuses mises au point afin d'améliorer la qualité et la pureté de l'échantillon. Les HS ainsi marqués et purifiés sont récoltés dans des quantités suffisamment importantes pour être analysés qualitativement et quantitativement par RMN afin d'en déterminer les caractéristiques structurales. En effet, l'enrichissement en ^{13}C permet l'étude de corrélation ^1H - ^{13}C dont les spectres obtenus sont plus informatifs et résolus que des spectres enregistrés en une seule dimension. L'établissement d'un protocole efficace, et validé avec un modèle de fibroblastes de souris embryonnaires (MEF) KO pour le gène codant la NDST-1 une enzyme importante dans la biosynthèse des HS, permet par la suite l'étude de l'évolution de structures des polysaccharides lors d'un phénomène de différenciation cellulaire.

Les résultats obtenus par RMN liquide permettent de mettre en place des expériences par RMN du solide visant l'analyse des HS directement au niveau des membranes plasmiques. L'étude des surfaces cellulaires par cette méthode sera présentée dans la dernière partie.

A. Mise en place d'un protocole de marquage et de purification d'HS entiers cellulaires pour leurs analyses RMN

Dans le but de mettre au point un protocole efficace de marquage et de purifications des HS, différents tests et vérifications préalables doivent être effectués.

Tout d'abord les cellules HeLa issues de cellules de cancer de col de l'utérus (d'Henrietta Lacks qui a donné son nom aux cellules) ont été choisies comme modèle cellulaire. Ce sont des cellules épithéliales facilement cultivables et exprimant une quantité importante de HSPG à leur surface (Takegawa et al. 2011) comme les syndécans et les glypicans (Sarrazin et al. 2011). Nous avons donc, dans un premier temps, vérifié la présence des différentes chaînes de GAG de type HS et CS au niveau des cellules par cytométrie en flux.

1. Vérification de l'expression en HS et CS des cellules HeLa par cytométrie en flux

Pour la vérification de la présence de GAG au niveau des surfaces membranaires, les cellules HeLa adhérentes cultivées dans des conditions standards sont détachées à 37°C de leur support par un chélateur d'ions bivalents (l'EDTA). Le fait de détacher les cellules avec une solution d'EDTA permet de préserver les HSPG et CSPG présents au niveau des membranes. Les cellules sont ensuite incubées à 4°C en présence d'anticorps (primaires) monoclonaux de type IgM dirigés spécifiquement contre les HS (mAb 10E4) ou les CS (mAb CS-56). Après cette première incubation et lavage des cellules, des anticorps secondaires anti-souris couplés à un fluorochrome FITC (*fluorescein isothiocyanate*) sont utilisés pour la détection par cytométrie en flux. L'utilisation d'un isotype IgM non spécifique constitue le contrôle négatif (Figure 68).

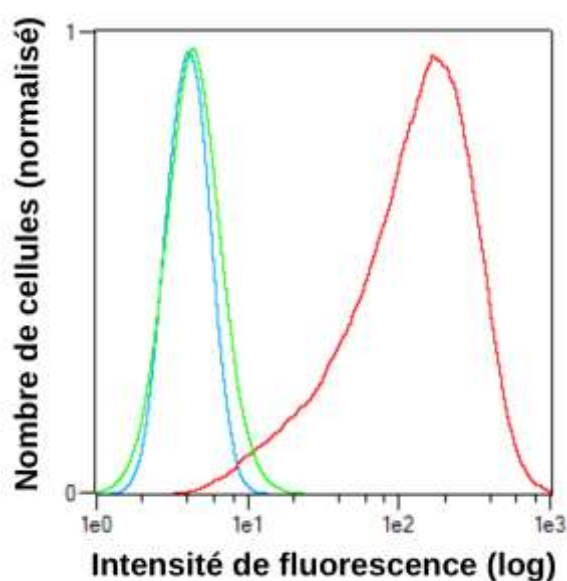


Figure 68 : Analyse par cytométrie en flux des cellules HeLa. Des anticorps spécifiques sont utilisés pour détecter les GAG à la surface des cellules. Détection des HS (rouge) et CS (vert), contrôle isotypique (bleu).

Les cellules incubées avec l'anticorps spécifique des HS montrent une intensité de fluorescence très importante (valeur de la médiane proche de 200) (courbe rouge). Cette intensité est comparée à celle du contrôle négatif (contrôle isotypique), correspondant à la fixation non-spécifique de l'anticorps secondaire sur les cellules (valeur de la médiane proche de 4). Cela montre que les cellules HeLa expriment effectivement à leur surface une quantité d'HS importante.

Pour la détection des CS, l'intensité de fluorescence obtenue (courbe verte) est comparable à celle de l'isotype contrôle, ce qui signifie que très peu de CS sont détectés.

Cette technique de cytométrie en flux avec marquage immunocytologique indirect (utilisation d'un anticorps primaire et d'un anticorps secondaire) ne permet pas une quantification précise des molécules (d'HS en l'occurrence). Cette intensité est également dépendante de l'affinité de chaque anticorps. Toutefois, l'analyse par cytométrie en flux a permis de vérifier la présence effective et importante des HS à la surface des cellules HeLa. Concernant les CS, une quantité très faible ou non-détectée a également été observée.

2. Conditions de culture et de marquage cellulaire

a) Conditions de culture

La culture des cellules HeLa en laboratoire dans des conditions appropriées est assez classique. Toutefois, bien que des antibiotiques (généralement pénicilline et streptomycine) soient généralement ajoutés au milieu de culture afin de prévenir les contaminations bactériennes, il a été observé que l'absence de ceux-ci permettait d'augmenter la vitesse de croissance des cellules HeLa, c'est pourquoi le milieu n'a pas été supplémenté en antibiotiques. De même, le milieu de culture utilisé est sans pyruvate afin d'habituer la cellule à produire son propre pyruvate à partir du glucose contenu dans le milieu. Mise à part ces spécificités, le milieu de culture est un milieu DMEM (*Dulbecco's modified eagle medium*) classique, additionné de sérum de veau fœtal et de glutamine pour la croissance et la prolifération cellulaire.

Puisque les HS sont destinés à être analysés par RMN et que cette technique est peu sensible, une quantité importante d'HS est nécessaire, et par conséquent une quantité importante de cellules les exprimant doit être cultivée. Dans ce but, des tests d'adaptation des cellules HeLa adhérentes en condition de non-adhérence ont été réalisés ce qui permet d'augmenter la concentration cellulaire, c'est-à-dire d'augmenter la quantité de cellules HeLa et donc d'HS récoltés pour un même volume de culture. Pour cela, les cellules sont initialement décollées de leur support par digestion à la trypsine afin de cliver les sites d'ancrage. Les cellules en solution sont ensuite placées dans un flacon *spinner* sous agitation magnétique constante et ménagée afin de limiter la mort cellulaire et de prévenir leur sédimentation. Les cellules sont diluées régulièrement afin de sélectionner les cellules survivant à leur nouvelle condition de culture. L'adaptation des cellules est suivie par un test de viabilité avec marquage au bleu de trypan permettant d'estimer le taux de mortalité.

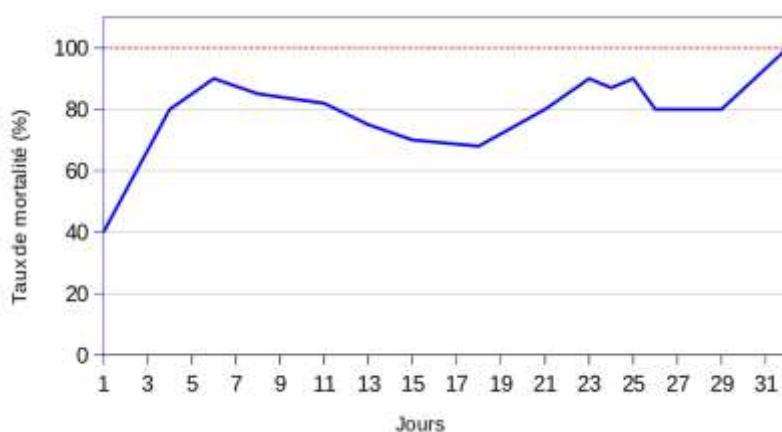


Figure 69 : Taux de mortalité des cellules HeLa adhérentes cultivées en spinner en condition de non-adhérence.

Au cours des premiers jours de culture (1 à 6) dans ces conditions, la mortalité cellulaire est progressive et très importante au 6e jour en atteignant un taux de 90% (Figure 69). La mortalité cellulaire diminue mais reste relativement stable (autour de 80%) jusqu'au 29e jour. Au-delà, 100% des cellules sont mortes (32e jour). Ce test montre qu'il est difficile d'adapter

les cellules HeLa croissant sur un support vers des conditions de culture en solution. Nous sommes donc restés dans des conditions de culture d'adhérence classique dans le milieu précédemment décrit.

Ains, les cellules HeLa adhérentes choisies comme modèle cellulaire, expriment une quantité d'HS importante à leur surface et cela a été vérifié par cytométrie en flux. Le but étant de réaliser des analyses qualitatives et quantitatives des HS par RMN, il est nécessaire de marquer les molécules d'intérêt avec l'isotope 13 du carbone (noté ^{13}C), isotope observable par RMN.

b) Détermination des conditions de marquage

Les HS sont constitués d'enchaînements d'hydrates de carbone. Ils sont majoritairement composés de noyaux d'hydrogène et de carbone. L'étude de ces noyaux permet par RMN de sonder au niveau de toute la chaîne d'HS l'environnement chimique des noyaux. Toutefois, les isotopes observables par RMN ne sont pas forcément les plus abondants à l'état naturel. Bien que le noyau d'hydrogène (^1H) composé d'un seul proton et observable par RMN représente plus de 99,9%, le noyau de carbone le plus abondant (^{12}C) n'est pas observable. Seul le ^{13}C (nucléairement stable) possède des propriétés physiques lui permettant d'être visible par RMN. Cependant, il ne représente que 1,1% à l'état naturel, c'est-à-dire qu'environ un noyau de carbone sur cent sera observé en RMN. Afin d'augmenter cette quantité en isotope et donc la quantité de signal, des techniques d'enrichissement ou dit "de marquage" doivent être mises en place.

Afin d'enrichir les chaînes d'HS en isotopes observables par RMN, le ^{13}C -glucose a été choisi pour supplémer le milieu de culture dépourvu de glucose. La présence de ^{13}C -glucose dans le milieu de culture va être la principale source carbonée pour les cellules et va entrer dans le métabolisme des HS car le glucose est le précurseur commun des GlcA et GlcNAc (Figure 17). L'utilisation de ^{13}C -glucose permet d'avoir le meilleur ratio efficacité de marquage/coût de production. L'inconvénient de l'utilisation de glucose marqué est son utilisation centrale dans la cellule pour la formation, notamment, d'autres éléments osidiques comme les riboses et désoxyriboses des ARN/ADN. Les différentes formes glycosylées vont donc être également marquées à l'isotope 13 du carbone et donc être observables par RMN, pouvant encombrer les spectres et rendre difficile les analyses RMN.

Le temps de marquage nécessaire pour que les cellules métabolisent efficacement le ^{13}C -glucose, le transforment et l'intègrent dans la synthèse des chaînes de GAG a été déterminé. Deux temps de marquage ont été testés, un temps "long" de 7 jours et un temps "intermédiaire" de 3 à 4 jours. Après purification et analyse par RMN des HS, dont les techniques seront décrites par la suite, il n'a pas été observé de différence significative. La principale différence entre les deux temps de marquage est que le temps de marquage dit long nécessite un renouvellement de milieu contrairement au temps de marquage intermédiaire. Le temps intermédiaire de marquage a donc été choisi car suffisant et nécessitant moins de ^{13}C -glucose.

Une fois marquées, les chaînes d'HS doivent être extraites des membranes cellulaires et purifiées pour leur analyse par RMN liquide.

3. Extraction, purification et analyse RMN des HS

a) Digestion des surfaces cellulaires

Afin de récupérer efficacement les chaînes d'HS à partir de surfaces cellulaires marquées, trois protéases citées dans la littérature pour la purification des GAG ont été testées : la trypsine (Miettinen et al. 1994), la papaïne (Lindahl et al. 1995) et la pronase (Pomin et al. 2010).

(1) Digestion des surfaces cellulaires à la trypsine

La trypsine est une endoprotéase qui permet de cliver les protéines dans leur partie C-terminale au niveau des résidus basiques (arginine et lysine) à 37°C et à pH physiologique. Les 150 millions de cellules marquées sont traitées avec la trypsine pendant 40mn. Après digestion, les cellules et les digestats des surfaces cellulaires sont récupérés puis centrifugés afin de séparer les cellules des chaînes d'HS en solution. Les cellules traitées sont analysées par cytométrie en flux après incubation avec des anticorps de reconnaissance spécifique des GAG (Figure 70).

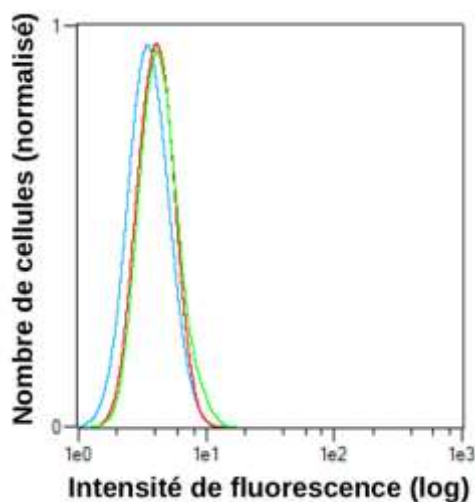


Figure 70 : Analyse par cytométrie en flux des cellules HeLa après digestion à la trypsine. Des anticorps spécifiques sont utilisés pour détecter les GAG à la surface des cellules. Détection des HS (rouge) et CS (vert), contrôle isotypique (bleu).

Les cellules incubées en présence de l'anticorps primaire spécifique des HS (courbe rouge) et des CS (courbe verte) montrent des intensités de fluorescence équivalentes à celles des cellules marquées avec l'isotype contrôle. Contrairement à ce qui est obtenu pour les HS en absence de traitement à la trypsine (Figure 68). Cette diminution drastique de l'intensité de fluorescence montre que le traitement à la trypsine permet de couper une grande majorité des PG cellulaires.

(2) Digestion des surfaces cellulaires à la papaïne

Une seconde enzyme à spectre large de digestion, la papaïne, a également été utilisée pour digérer les surfaces cellulaires. En effet, la papaïne clive les protéines en position C-terminale des acides aminés basiques (lysine et arginine) avec un acide aminé hydrophobe N-terminal. La température optimale d'activité de l'endoprotéase n'est pas compatible avec les conditions de cultures cellulaires. En effet, la température d'activité de la papaïne est de 65°C et requiert un temps de digestion long (16h). Ces conditions ne sont pas compatibles avec la survie des cellules, ce qui rend impossible les analyses par cytométrie en flux des surfaces cellulaires qui ont perdu leur intégrité.

(3) Digestion des surfaces cellulaires à la pronase

Enfin une troisième et dernière enzyme a été utilisée, il s'agit de la pronase qui est en réalité un cocktail de protéases (contenant des sérine-protéases, métalloprotéases, carboxypeptidases et aminopeptidases) permettant une large digestion des protéines. Les cellules sont marquées puis incubées en présence de ces enzymes à 37°C pendant 30mn. Les conditions de digestion sont plus ménagées comparées aux conditions de digestion à la papaïne. Le taux de survie des cellules avant et après digestion est similaire et supérieur à 90%. Toutefois l'analyse par cytométrie en flux des surfaces cellulaires digérées à la pronase montre la présence d'HS après traitement (Figure 71). Le cocktail d'enzyme n'est donc pas suffisamment efficace comparé au traitement trypsin des surfaces cellulaires qui permet de solubiliser tous les HS des surfaces cellulaires.

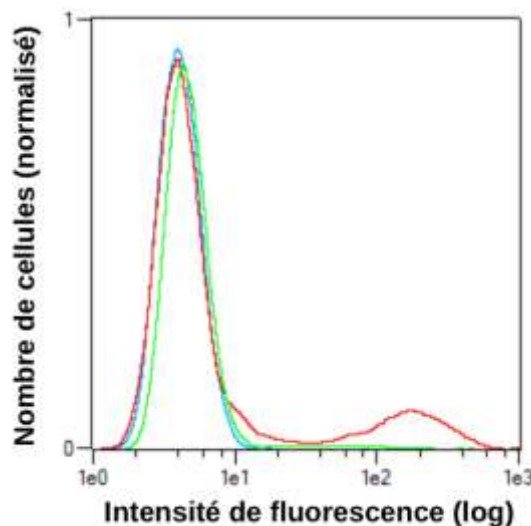


Figure 71 : Analyse par cytométrie en flux des cellules HeLa après digestion à la pronase. Des anticorps spécifiques sont utilisés pour détecter les GAG à la surface des cellules. Détection des HS (rouge) et CS (vert), contrôle isotypique (bleu).

Les digestats, issus des trois digestions des surfaces cellulaires, contenant les HS sont ensuite purifiés par chromatographie échangeuse d'anions puis la présence des HS est observée par RMN.

b) Purification et analyse RMN des digestats cellulaires

Dans un premier temps les échantillons sont purifiés par chromatographie échangeuse d'anions du fait de la charge nette négative des chaînes d'HS. Toutes les molécules d'interaction avec la matrice DEAE (DiEthylAminoEthyl) chargée positivement sont éluées à une concentration de 1,5M NaCl qui va casser les interactions électrostatiques. L'élution par chromatographie est suivie par l'absorbance des protéines à 280nm et 215nm et des acides nucléiques à 260nm. Toutefois, aucune longueur d'onde n'étant spécifique des HS entiers, seule l'analyse RMN permet de vérifier la présence des HS.

(1) Purification et analyse RMN de la digestion à la trypsine des surfaces cellulaires

Le digestat des surfaces cellulaires après traitement à la trypsine est chargé sur une matrice DEAE. Dans un premier temps, les molécules de très faible interaction ou n'interagissant pas avec la matrice sont éliminées par rinçage de la colonne avec du PBS (pH 6,8) à une faible concentration en sel (150mM NaCl). Les HS d'intérêt, polysulfatés et donc interagissant fortement avec la DEAE, sont élués avec une concentration en sel de 1,5M NaCl (Figure 72).

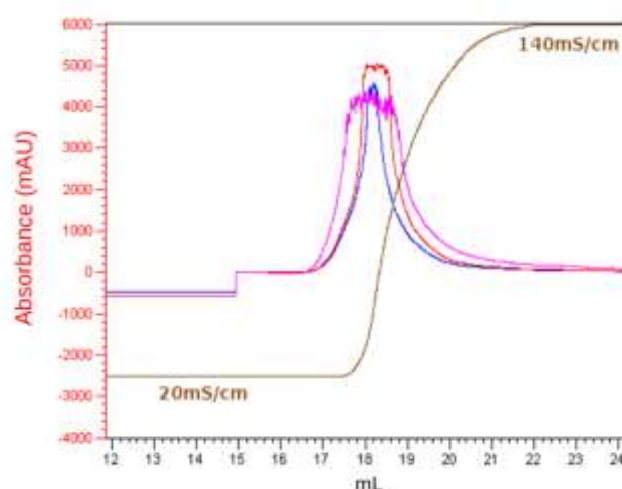


Figure 72 : Purification après digestion à la trypsine de cellules HeLa marquées. Le chromatographe montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat.

L'absorbance de la fraction retenue (17-21mL) est saturée (>4000mAU) pour les différentes longueurs d'onde observées à 215nm, 260nm et 280nm. Cela indique une quantité relativement importante en contaminants protéiques et nucléiques.

Cette fraction d'élution est ensuite dessalée, lyophilisée et solubilisée dans de l'eau lourde ($^2\text{H}_2\text{O}$ ou D_2O) pour son analyse par RMN par des techniques ^1H - ^{13}C HSQC permettant de déterminer les corrélations entre ^1H et ^{13}C directement liés. Les largeurs spectrales définies entre 3ppm et 6,5ppm en ^1H et entre 50ppm et 110ppm en ^{13}C permettent d'observer les corrélations attendues pour les sucres (Figure 73A). Les contaminants protéiques ne sont donc pas observés. Le spectre obtenu montre la présence de différents sucres, et notamment des HS (Figure 73A) par comparaison avec le spectre d'HS commercial (Figure 73B). Toutefois, les

différentes corrélations observées ne correspondent pas toutes aux HS, et certaines sont caractéristiques des acides nucléiques (Figure 73A). En effet, les acides nucléiques sont en partie faits d'oses (désoxyriboses et riboses) qui sont synthétisés par la cellule à partir du ^{13}C -glucose ajouté au milieu de culture. Les déplacements chimiques de ces oses sont proches de ceux des HS, notamment dans les corrélations en-dehors des carbones anomériques.

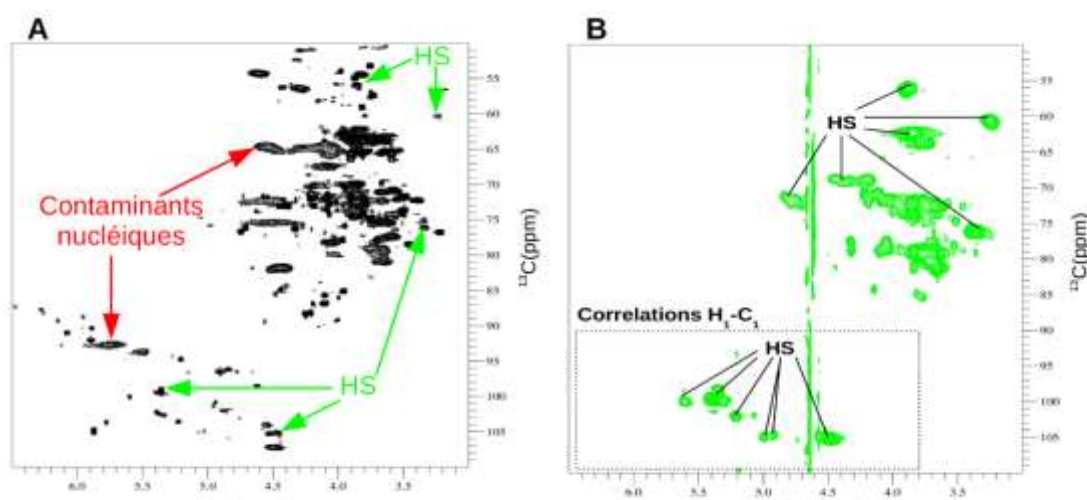


Figure 73 : Analyse par RMN après digestion à la trypsine de cellules HeLa marquées. Le spectre ^{13}C -HSQC (A) est partiellement attribué pour les HS (flèches vertes) et pour les acides nucléiques (flèches rouges) et comparé à un spectre ^{13}C -HSQC d'HS commerciaux (B).

En résumé, la purification et l'analyse spectrale permettent de mettre en évidence la présence claire d'HS marqués au sein des échantillons issus des digestions des cellules HeLa par la trypsine. Toutefois, cette première étape de purification montre également la présence de contaminants protéiques, et nucléiques. L'analyse spectrale des HS dans ces conditions ne peut pas être quantitative du fait de la présence des riboses/désoxyriboses dont les déplacements chimiques sont proches de ceux des HS et pas suffisamment résolus.

(2) Purification et analyse RMN après digestion à la papaïne des surfaces cellulaires

Le digestat des cellules traitées à la papaïne est purifié par chromatographie échangeuse d'anions. Ce dernier est élué par un pas en concentration de sel (NaCl) de 0,15 à 1,5M pour casser les interactions avec la matrice DEAE (Figure 74A). L'absorbance de la fraction récoltée (12-20mL) est saturée ($>4000\text{mAU}$) pour les 3 longueurs d'onde observées (215nm, 260nm et 280nm), indiquant une quantité importante de contaminants. L'analyse plus quantitative de l'échantillon en contaminants nucléiques est réalisée par son absorbance à 260nm au Nanodrop qui permet une détermination de la densité optique (DO) plus précise (Figure 74B). La DO obtenue à 260nm est de 270, ce qui est à la limite supérieure de l'intervalle de confiance de l'appareil. Sachant qu'une unité de DO correspond à une concentration de $50\mu\text{g.mL}^{-1}$ d'ADN double brin et que le volume de fraction est de 8mL, une centaine de milligrammes d'acides nucléiques est récoltée ($270\text{DO} \times 0,05\text{mg.mL}^{-1} \times 8\text{mL} = 108\text{mg}$).

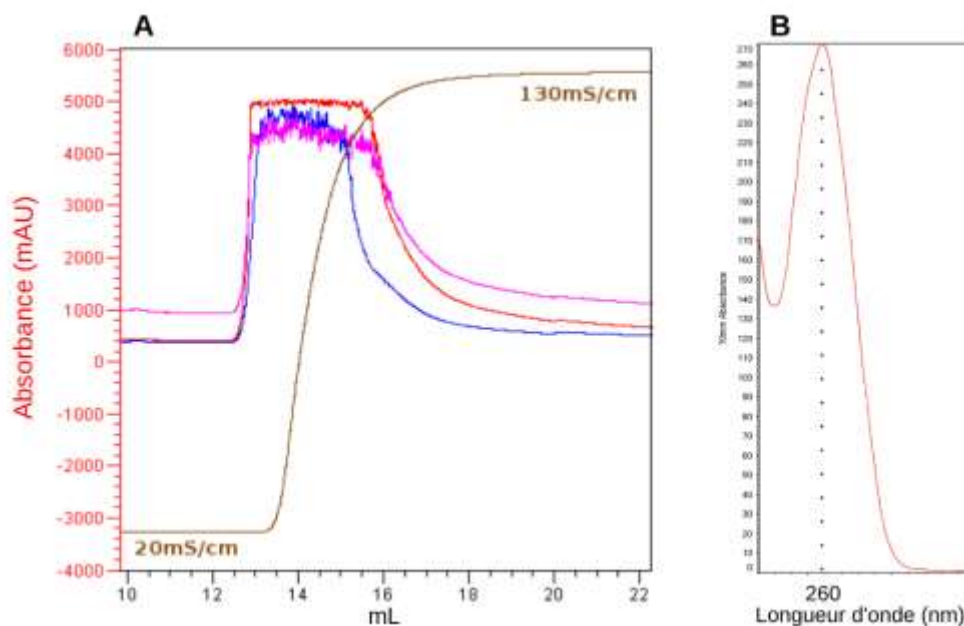


Figure 74 : Purification et analyses UV après digestion à la papaïne de cellules HeLa marquées. Le chromatographe (A) montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat. Absorbance UV de l'échantillon (B).

La fraction est également analysée par RMN. Les corrélations ^1H - ^{13}C montrent effectivement en quantité très importante les acides nucléiques à travers les corrélations des riboses et désoxyriboses fléchées en rouge (Figure 75A). Toutefois, malgré le recouvrement des pics dû à la présence importante en ADN/ARN marqués dans l'échantillon, des corrélations appartenant aux HS (fléchées en vert) sont également observées par comparaison avec un spectre d'HS commercial (Figure 75B), mais avec des intensités plus faibles comparées aux signaux des HS obtenus à partir d'HS de cellules HeLa digérées à la trypsine (Figure 73A).

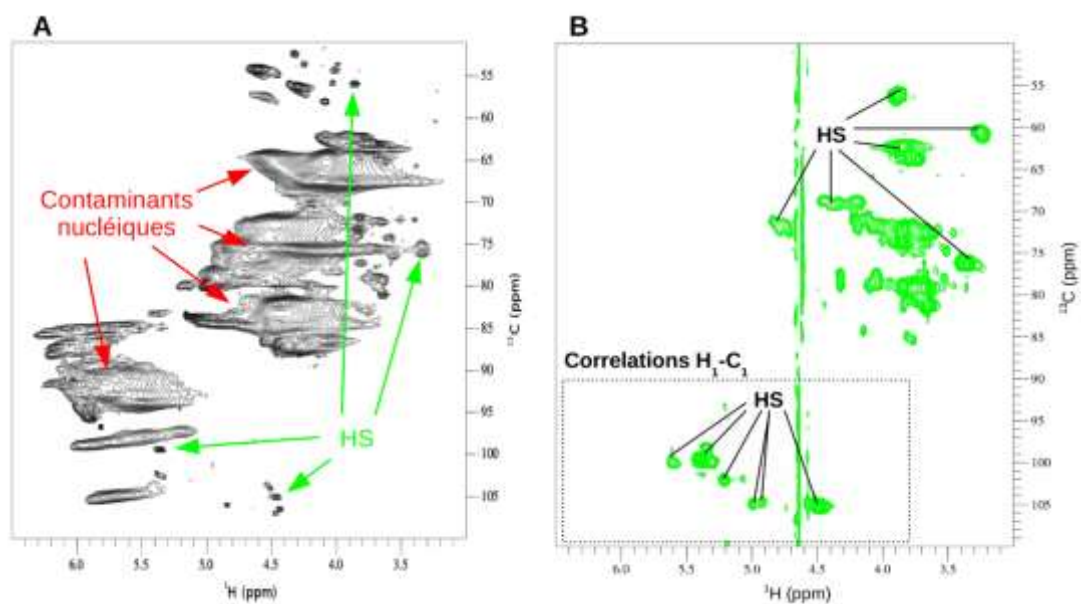


Figure 74 : Analyse par RMN après digestion à la papaïne de cellules HeLa marquées. Le spectre ^{13}C -HSQC (A) est partiellement attribué pour les HS (flèches vertes) et pour les acides nucléiques (flèches rouges) et comparé à un spectre ^{13}C -HSQC d'HS commerciaux (B).

La digestion des surfaces de cellules HeLa à la papaïne se fait dans des conditions létales pour les cellules qui larguent toute ou partie de leur matériel génétique. Ces contaminants nucléiques, du fait de leur nature similaire aux HS (longs polymères anioniques), vont également être purifiés par chromatographie et être observés par RMN. Les HS et les ADN/ARN vont être en compétition pour les sites de fixation à la matrice DEAE, ce qui peut diminuer le rendement d'HS obtenus. Les analyses RMN avec de telles concentrations en acides nucléiques ne sont pas possibles à cause du recouvrement spectral entre les riboses/désoxyriboses et glucosamines et acides uroniques. La digestion des surfaces cellulaires par la trypsine est préférée à la digestion à la papaïne car elle permet d'avoir un échantillon brut moins contaminé, facilitant les étapes de purification et d'analyse.

(3) Purification et analyse RMN après digestion au cocktail pronase des surfaces cellulaires

Comme vu précédemment, le digestat des cellules par la pronase est purifié par chromatographie échangeuse d'anions en pas de sel (0,15-1,5M NaCl). La fraction récoltée est de 5mL (16,5-21,5mL) (Figure 76) et est donc inférieure à la fraction obtenue après purification des digestats des cellules digérées à la papaïne, mais comparable au volume récolté pour les digestats à la trypsine. Les absorbances des protéines à 215nm et 280nm ainsi que des acides nucléiques à 260nm sont saturées (>4000mAU) et permettent d'estimer la présence de ces éléments contaminants dans des quantités relativement importantes.

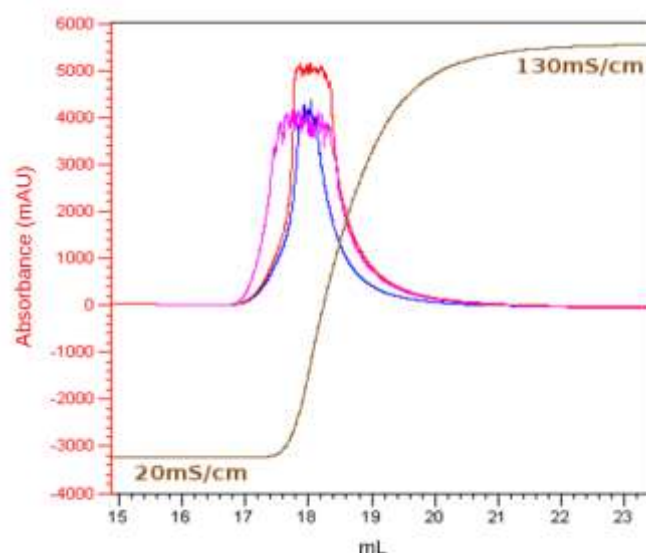


Figure 76 : Purification après digestion à la pronase de cellules HeLa marquées. Le chromatographe montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat.

L'analyse par RMN permet de confirmer la présence en quantité importante des acides nucléiques (Figure 77A) dont les corrélations sont indiquées par des flèches rouges. L'attribution du spectre RMN ^1H - ^{13}C HSQC montre également la présence significative d'HS au sein de l'échantillon dont les pics de corrélation de ces espèces sont indiqués par des flèches vertes.

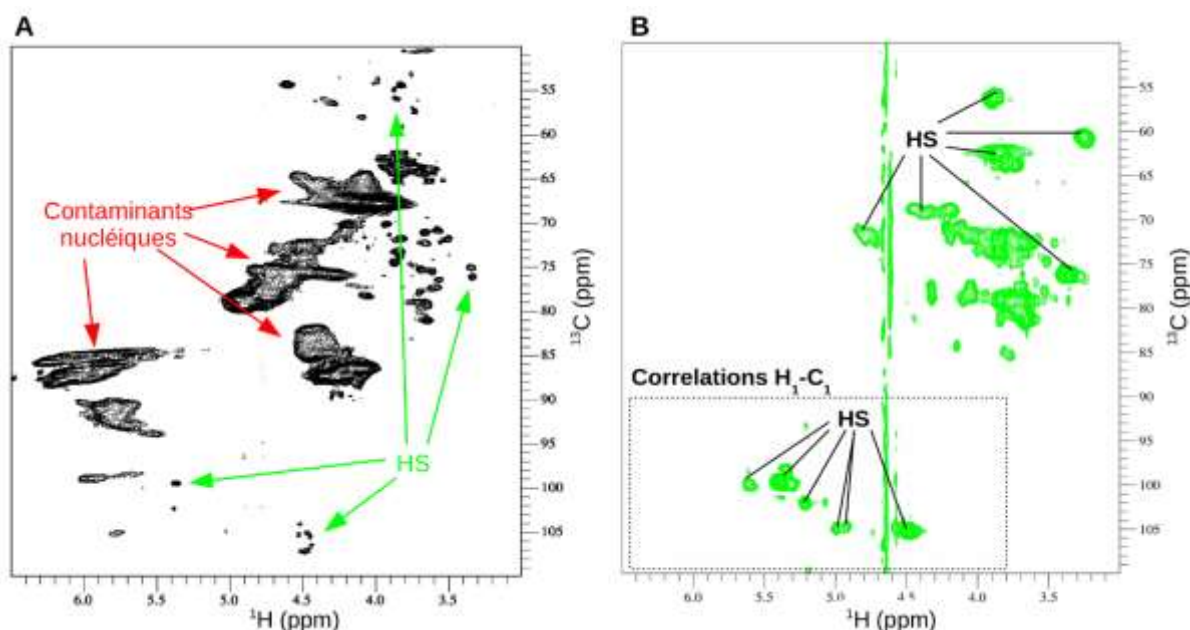


Figure 77 : Analyse par RMN après digestion à la pronase de cellules HeLa marquées. Le spectre ^{13}C -HSQC (A) est partiellement attribué pour les HS (flèches vertes) et pour les acides nucléiques (flèches rouges) et comparé à un spectre ^{13}C -HSQC d'HS commerciaux (B).

Les signaux obtenus par RMN pour cette préparation d'HS semblent être moins intenses que ceux obtenus à partir des HS extraits de cellules HeLa digérées à la trypsine. Les analyses par cytométrie en flux – après digestion des surfaces cellulaires au cocktail pronase ont permis de montrer que quelques chaînes d'HS sont encore présentes après digestion (Figure 71) contrairement au traitement à la trypsine qui permet une digestion totale (Figure 70).

En somme, les résultats précédents montrent que le traitement à la papaïne à cause des conditions de digestion n'est pas idéal car il provoque la mort des cellules et le relargage de leur contenu génétique, polluant de ce fait les étapes de purification et d'analyse. La digestion "pronase" quant à elle permet de limiter cette contamination aux ADN/ARN, mais l'analyse RMN ne montre pas une quantité importante en HS comparée à la digestion des cellules par la trypsine pour une même quantité cellulaire. Cette dernière permet en effet de limiter la présence de contaminants dans la fraction étudiée, et montre une digestion des surfaces cellulaires optimale pour la solubilisation des chaînes d'HS par clivage des PG.

Pour la suite de la purification, les échantillons obtenus à partir des digestions des cellules HeLa par la trypsine et purifiés par chromatographie échangeuse d'anions doivent être digérés de manière plus approfondie, et principalement en digérant les acides nucléiques qui gênent les analyses spectrales.

c) Digestion des principaux contaminants, les acides nucléiques

Les premières purifications présentées précédemment ont mis en évidence la présence de sucres marqués qui peuvent être analysés par RMN dans une certaine mesure. Parmi eux, se trouvent les HS d'intérêt pour notre étude, mais également les oses d'acides nucléiques. Du fait de leur présence, les études par RMN ne permettent pas de caractériser finement les HS

car les déplacements chimiques des différentes espèces ne sont pas toujours suffisamment résolus. Les fractions obtenues après cette première étape de digestion protéique et de chromatographie échangeuse d'anions doivent encore être débarrassées des contaminants ADN/ARN. Pour cela, deux enzymes, la benzonase et la DNase I, décrites dans la littérature pour la purification des HS (Lindahl et al. 1995) (Habuchi et al. 2007), ont été testées afin de déterminer leur efficacité dans nos conditions expérimentales.

(1) Digestion à la benzonase, purification et analyse RMN d'HS cellulaires

Des tests de digestion ont, en premier lieu, été réalisés à la benzonase sur les fractions récoltées après une première étape de purification par chromatographie des HS issus de cellules digérées à la trypsine. La benzonase est une endonucléase à spectre de digestion large qui dégrade les ARN et les ADN. Les échantillons sont digérés à 37°C pendant 21h puis chargés sur une colonne de chromatographie échangeuse d'anions pour leur purification par la même méthode chromatographique (Figure 78). La fraction d'élution qui précédemment avait une absorbance saturante aux différentes longueurs d'onde (215nm, 260nm et 280nm) a maintenant des absorbances quantifiables (1000mAU à 280nm, 2200mAU à 260nm et plus de 2400mAU à 215nm). Cette baisse générale de l'absorbance montre qu'il y a eu effectivement une dégradation des acides nucléiques après traitement à la benzonase mais également une dégradation au cours du temps des contaminants protéiques.

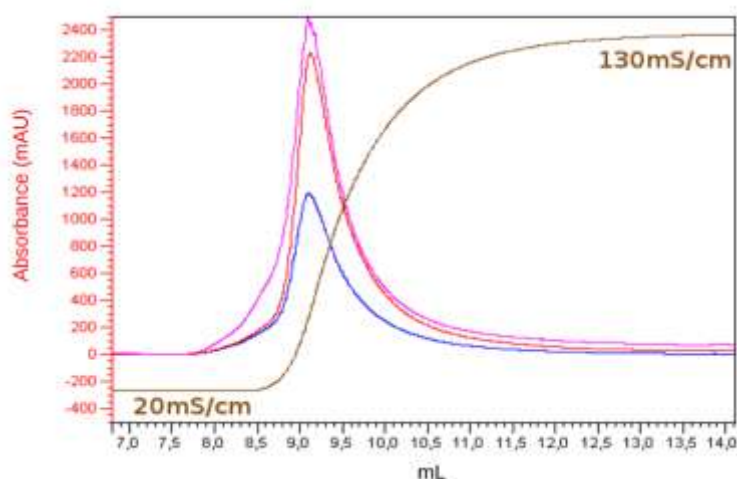


Figure 78 : Purification après digestion à la benzonase d'HS extraits de cellules HeLa marquées digérées à la trypsine. Le chromatographe montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat.

La fraction d'élution (8-11mL) est ensuite analysée par RMN (Figure 79). L'attribution partielle de l'échantillon permet d'observer toujours les pics attendus pour les HS (spectre noir) observés avant digestion à la nucléase (spectre cyan). Cependant, les corrélations des riboses et désoxyriboses sont toujours observées dans des quantités importantes dans cet échantillon, et ce malgré une absorbance à 260nm plus faible. La benzonase apparaît donc comme une enzyme ne permettant pas la digestion totale des acides nucléiques qui sont toujours présents dans des échantillons complexes issus de digestion de cellules. C'est

pourquoi, une deuxième nucléase a également été testée.

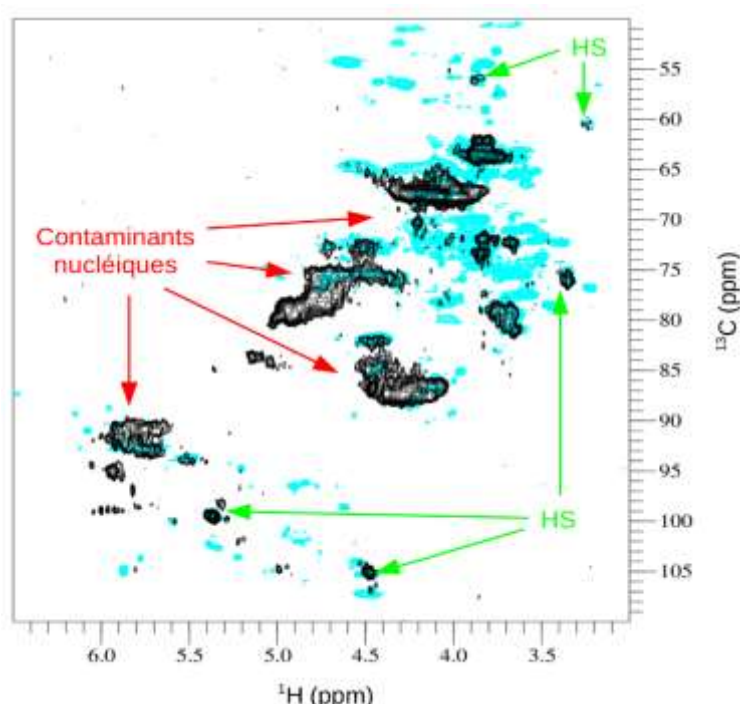


Figure 79 : Analyse par RMN de l'extrait d'HS de cellules HeLa marquées avant (spectre cyan) et après (spectre noir) digestion à la benzonase. Les spectres ^{13}C -HSQC sont partiellement attribués pour les HS (flèches vertes) et pour les acides nucléiques (flèches rouges).

(2) Digestion à la DNase I, purification et analyse RMN d'HS cellulaires

L'utilisation de DNase I a également été décrite dans la littérature (Habuchi et al. 2007). Les fractions d'élution récoltées après digestion des surfaces cellulaires à la trypsine et purifiées sur colonne échangeuse d'anions sont digérées dans des conditions similaires à la benzonase, à 37°C mais dans des temps plus longs (24h) puis chargées sur une colonne de DEAE. Le chromatographe dans sa partie d'élution en 1,5M NaCl est présenté Figure 80. La fraction (6-8mL) a une absorbance à 280nm inférieure à 500mAU indiquant une quantité réduite de contaminants de nature protéique. Une forte diminution de l'absorbance est également observée à 260nm qui n'est plus que de 700mAU après traitement à la DNase I contre une absorbance qui saturait (>4000mAU) avant traitement à cette nucléase (Figure 72).

L'analyse spectroscopique 2D de corrélation ^1H - ^{13}C par RMN montre toujours la présence des corrélations attendues pour les HS (Figure 81), mais les corrélations des oses des acides nucléiques (désignées par les flèches rouges) sont nettement moins intenses dans cette analyse comparée à la fraction digérée à la benzonase (Figure 79). Dans des conditions spectroscopiques similaires, la baisse de cette intensité est liée à une baisse de la quantité d'acides nucléiques au sein de l'échantillon, ce qui est aussi vérifié par la comparaison des absorbances UV à 260nm. Un échantillon avec un degré de pureté similaire permet d'envisager des études structurales plus fines par RMN, comme des analyses quantitatives et une détermination précise par attribution des différentes unités disaccharidiques composant les

HS issus de surfaces de cellules HeLa.

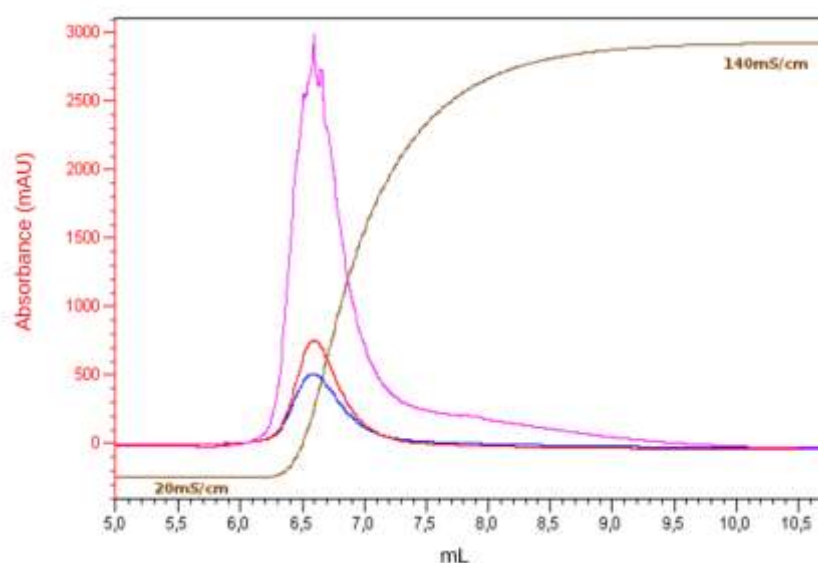


Figure 80 : Purification après digestion à la DNase I d'HS extraits de cellules HeLa marquées digérées à la trypsine. Le chromatographe montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat.

En conclusion, pour une digestion complète des acides nucléiques présents dans les extraits issus des digestions cellulaires, la benzonase se révèle relativement peu efficace. En effet, la digestion n'est pas totale dans les quantités d'enzyme et temps expérimentaux mis en place. La DNase I montre cependant une meilleure efficacité d'après le spectre de digestion. Le traitement à la DNase I permet d'obtenir un échantillon suffisamment pur pour des analyses RMN plus poussées. Par conséquent, l'utilisation de la DNase I a finalement été préférée à la benzonase. De plus les conditions de digestion de la DNase I (pH d'activité optimal de l'enzyme

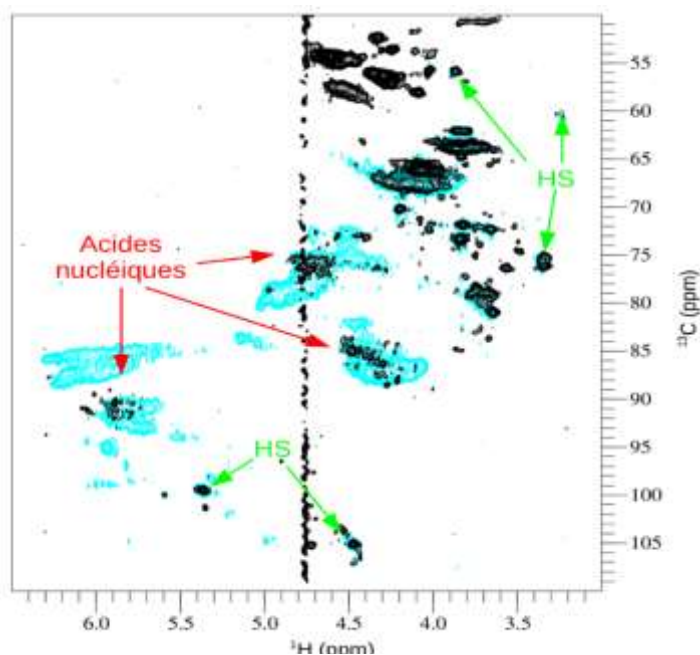


Figure 81 : Analyse par RMN de l'extrait d'HS de cellules HeLa marquées avant (spectre cyan) et après (spectre noir) digestion à la DNase I. Les spectres ^{13}C -HSQC sont partiellement attribués pour les HS (flèches vertes) et pour les acides nucléiques (flèches rouges).

plus proche du pH du tampon de purification, présence de 0,15M NaCl dans le tampon d'enzyme comme dans le tampon de purification...) permettent une purification plus simple de l'échantillon.

Ainsi, un protocole de purification efficace a pu être élaboré, permettant d'envisager une analyse par RMN qualitative et quantitative. Cependant, des améliorations à ce protocole peuvent être apportées dans le but d'augmenter la quantité d'HS récoltés, d'optimiser les étapes de chromatographie et de réduire des pertes de matériel dues aux nombreuses manipulations.

4. Amélioration des conditions d'extraction et de purification des HS cellulaires

Afin d'obtenir une quantité de 200 millions de cellules, 10 flasques de culture de type T175 sont nécessaires. En plus de l'espace important occupé au niveau de l'incubateur CO₂, ces 10 flasques nécessitent beaucoup de manipulation. C'est pourquoi des tests sur différents types de flasques, avec une surface totale identique mais occupant un espace plus réduit et permettant moins de manipulations, ont été réalisés.

a) Choix des flasques de culture

Du fait de la faible sensibilité de la technique RMN, une production importante d'HS doit être réalisée afin d'avoir suffisamment de matériel pour leur analyse fine. Nous avons donc testé des flasques adaptées à la culture d'un grand nombre de cellules dans un volume restreint. Deux systèmes différents ont été testés pour l'extraction d'HS à partir de 200 millions de cellules.

(1) Système HYPERflask™ (Corning Life Science®)

Ce système correspond à une surface de culture de 1750cm² fortement compactée (voir Figure 65). Après marquage des cellules au ¹³C-glucose, les PG des surfaces de ces dernières doivent être digérés afin de libérer les chaînes de GAG et permettre leur purification et analyse spectrale. La configuration de ces flasques est telle qu'il n'est pas possible de recouvrir toute la surface de culture avec un volume raisonnable de solution de trypsine. Une agitation manuelle est nécessaire afin de pouvoir digérer efficacement toutes les cellules contenues dans la flasque de culture. Ces conditions d'extraction assez drastiques sont délétères pour les cellules qui larguent en abondance leur matériel génétique. De plus, du fait de ces contraintes, il est difficile de réaliser un comptage des cellules marquées, toutefois nous pouvons considérer que le nombre de cellule reste proche de 200 millions de cellules. Après digestion des surfaces cellulaires et centrifugation, le surnageant doit ensuite être filtré (diamètre de 1,2µm) afin d'éliminer les éléments insolubles.

Après filtration, l'échantillon est purifié par chromatographie échangeuse d'anions (Figure 82A). Les interactions matrice-anions sont cassées par une forte concentration saline. Cette concentration en sel est suivie au travers de la conductivité (courbe marron) de la solution analysée en sortie de colonne et est donnée en milliSiemens par centimètre. Les

éléments élués sont détectés par leur absorbance UV à 260nm pour les acides nucléiques et à 215nm et 280nm pour les protéines. Les absorbances (données en unités arbitraires AU) de l'éluat (11-16mL) saturé au niveau du détecteur (>4000mAU) aux différentes longueurs d'ondes, indiquant la présence en quantité importante de protéines et d'acides nucléiques. La fraction d'éluat est ensuite analysée par RMN (Figure 82B).

Le spectre de corrélation ^1H - ^{13}C obtenu et zoomé dans une zone de déplacements chimiques particuliers attendus pour les HS montre effectivement la présence importante des sucres des ADN et ARN à 5,4-6,2ppm en ^1H et 90-95ppm en ^{13}C (encadré rouge) et corrèle avec les données d'absorbance UV à 260nm. L'analyse RMN de la fraction révèle également la présence de pics de corrélation appartenant aux héparanes sulfates (cercles verts).

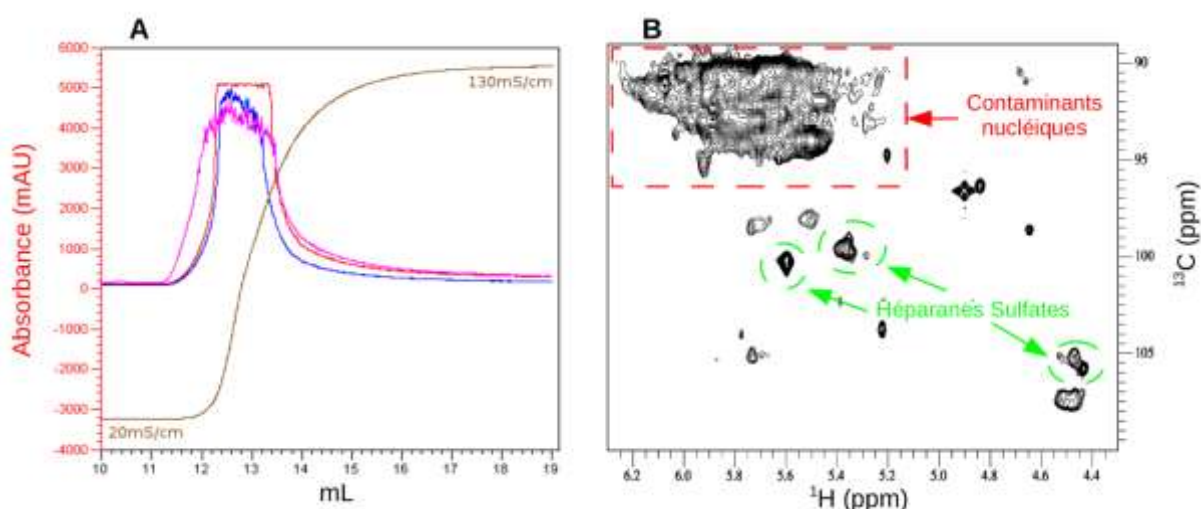


Figure 82 : Purification et analyse RMN d'une digestion à la trypsine de cellules HeLa cultivées en système HYPERflask™. Le chromatographe (A) montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat. Le spectre de RMN ^1H - ^{13}C HSQC (B) est partiellement attribué pour les HS (flèches et cercles verts) et pour les acides nucléiques (encadré rouge).

Ainsi l'extraction, la purification et la détermination par RMN de la digestion des surfaces de cellules HeLa cultivées dans le système HYPERflask™ rendent compte d'un degré de contamination aux protéines et aux acides nucléiques important. Ces derniers possédant également un ose ont un déplacement chimique ^1H - ^{13}C similaire aux HS et sont observés dans la même fenêtre spectrale, contrairement aux protéines, ce qui peut rendre difficile une analyse qualitative fine par la suite.

Un second système de flasque de culture cellulaire a également été testé.

(2) Système T600 (EDM Millipore)

Le système T600 correspond à une triple flasque (3 surfaces cellulaires de 200cm²) et est apparenté au système de culture cellulaire classique, mais avec un volume de boîte plus restreint et donc plus facile à manipuler. La surface de culture par boîte est de 600cm², ce qui est inférieur au système HYPERflask™. Pour une même quantité de cellules cultivées, trois T600 sont nécessaires, ce qui augmente le nombre de manipulations mais cela reste tout à fait acceptable.

Après marquage, les surfaces cellulaires sont digérées plus aisément, la solution enzymatique de trypsine pouvant recouvrir toutes les surfaces cellulaires et ne nécessitant pas d'agitation mécanique. La mort cellulaire est alors bien moins importante, et une simple étape de centrifugation permet de séparer le digestat des cellules. La solution contenant les GAG cellulaires est purifiée comme précédemment par chromatographie échangeuse d'anions (Figure 83A). Un *step* de sel permet de casser les interactions matrice-anions. La fraction d'élution (17-21mL) est saturée ($>4000\text{mAU}$) aux différentes longueurs d'onde observées (215nm, 260nm et 280nm).

L'analyse RMN ^1H - ^{13}C de l'échantillon obtenu (Figure 83B) montre clairement les corrélations attribuées aux HS (cercles verts). La présence de contaminants nucléiques est toujours observée (5,8-91ppm ^1H - ^{13}C) (encadré rouge), mais le ratio de l'intensité des pics des riboses/désoxyriboses par rapport à l'intensité des pics d'HS est nettement moindre, ce qui montre, malgré des absorbances saturantes à 260nm pour chaque chromatographie, que la quantité en acides nucléiques est plus faible après extraction d'une même quantité d'HS dans le système T600.

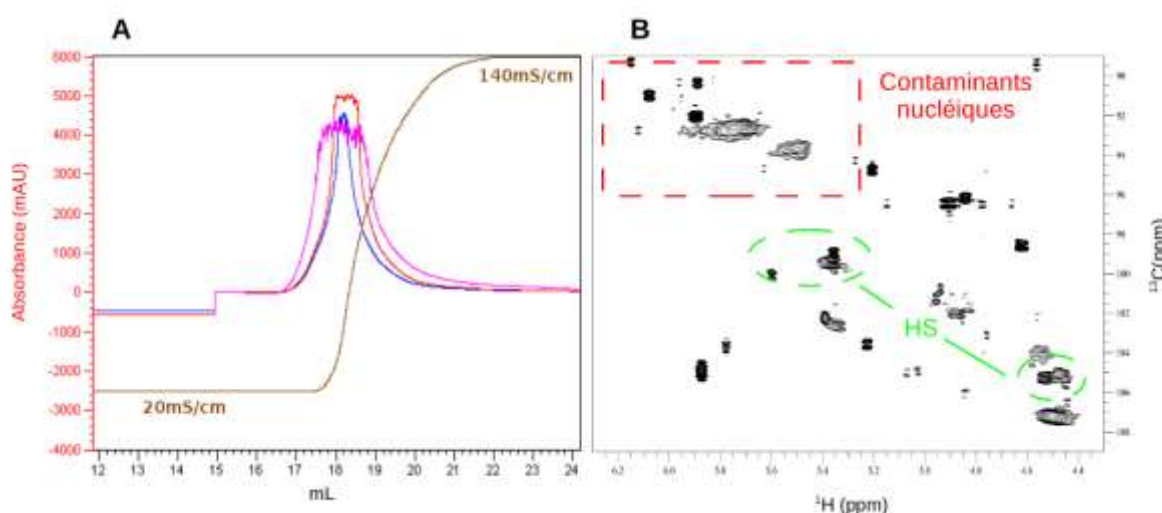


Figure 83 : Purification et analyse RMN d'une digestion à la trypsine de cellules HeLa cultivées en système T600. Le chromatographe (A) montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat. Le spectre de RMN ^1H - ^{13}C HSQC (B) est partiellement attribué pour les HS (traits et cercles verts) et pour les acides nucléiques (encadré rouge).

La flasque T600, bien que plus restreinte en surface, offre un avantage certain pour la digestion des surfaces cellulaires, contrairement au système HYPERflask™. La perte de surface de culture peut être aisément compensée par l'augmentation du nombre de T600. Les conditions de digestion des surfaces cellulaires étant plus faciles et moins délétères pour les cellules, cela permet d'obtenir un premier digestat brut moins contaminé, notamment aux acides nucléiques dont les déplacements chimiques de leur ose peuvent être problématiques pour l'analyse RMN. A la vue des nombreux avantages comparés au système HYPERflask™, les T600 ont été logiquement choisies pour les préparations cellulaires.

b) Optimisation de la méthode d'élution des HS par chromatographie échangeuse d'anions

Différentes techniques d'élution ont été testées dans le but de récupérer dans la fraction d'intérêt un maximum d'HS d'intérêt. Une première méthode a été d'éluer les molécules anioniques d'interaction avec un changement de concentration en NaCl non-progressif (de 0,15 à 1,5M NaCl) afin de casser toutes les interactions électrostatiques (faibles, intermédiaires et fortes) créées avec la matrice DEAE. Cette technique est dite "en pas" de sel, ou *step*. Une seconde méthode quant à elle dite en « gradient de sel » permet d'augmenter progressivement la concentration en sel et de détacher de la colonne les différentes molécules anioniques selon leur charge nette, les plus grosses étant éluées en dernier.

(1) Éluion en step

Un chromatographe type de cette méthode de purification est illustré Figure 84A. Lors de l'augmentation de concentration en NaCl, suivie par l'augmentation de la conductivité

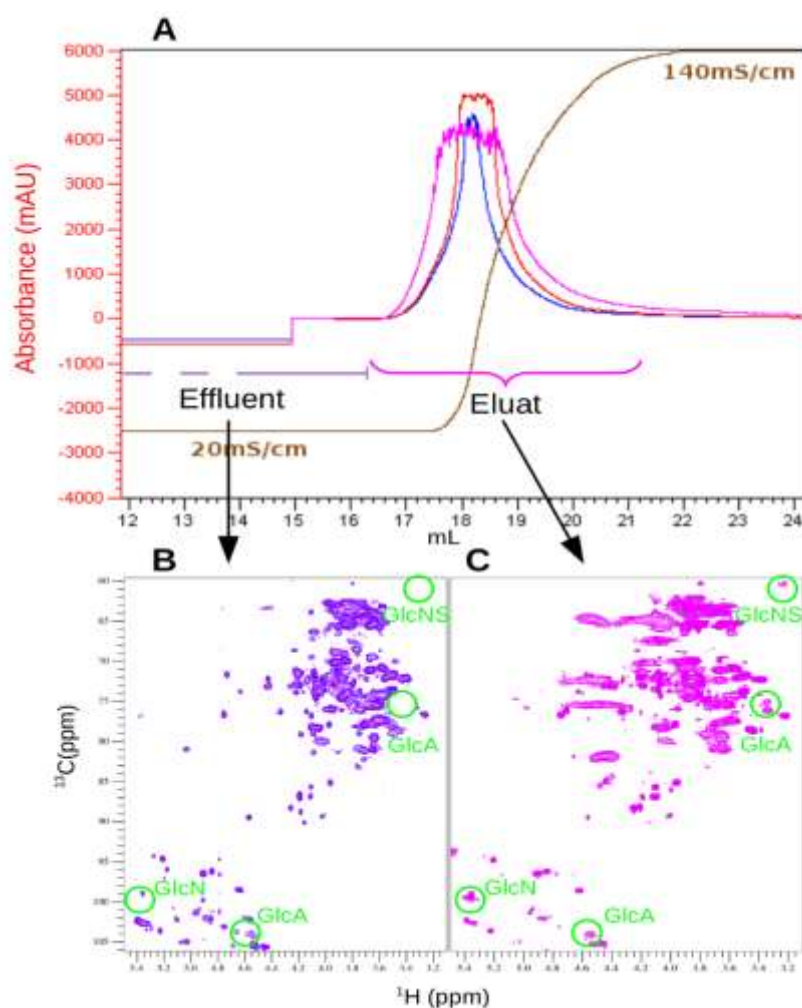


Figure 84 : Purification par palier de sel et analyse RMN d'une digestion à la trypsine de cellules HeLa. Le chromatographe (A) montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat. Les spectres de RMN ^1H - ^{13}C HSQC de l'effluent (B) et de l'éluat (C) sont partiellement attribués pour les HS (cercles verts).

(courbe marron), toutes les molécules anioniques absorbant à 215nm, 260nm et 280nm semblent être élués entre 16mL et 22mL. La fraction précédente (0-16mL) élue avec 150mM NaCl constitue le *flowthrough* ou effluent c'est-à-dire la fraction contenant des molécules n'interagissant pas ou très faiblement avec la matrice. Pour plus de lisibilité, le chromatographe est zoomée dans sa partie d'élution de 12 à 24mL. Une absorbance saturant aux différentes longueurs d'onde est observée dans les premiers millilitres de rinçage de la colonne (non-montré). Cet effluent est ensuite dessalé, lyophilisé et solubilisé dans de l'eau lourde (D₂O) pour son analyse par RMN bidimensionnelle HSQC de corrélation ¹H-¹³C (Figure 84B).

Les déplacements chimiques attendus pour les différents noyaux des HS sont encerclés en vert sur le spectre RMN. Dans cette fraction, toutes les corrélations ne sont pas observées, notamment les corrélations ¹H₂-¹³C₂ des GlcA (3,36-76,5ppm ¹H-¹³C) des glucosamines *N*-sulfatées (3,25-61ppm ¹H-¹³C). La fraction *flowthrough* de cette purification ne contient donc pas les HS d'intérêt.

En revanche, l'analyse spectroscopique de RMN de l'éluat à forte concentration en NaCl (16-22mL) révèle toutes les corrélations attendues pour les HS encerclées en vert (Figure 84C) pour les acides uroniques et les glucosamines. Le spectre montre également une certaine quantité en composés sucrés n'appartenant pas aux GAG.

(2) Éluion en gradient de sel

Une seconde méthode de purification testée a été d'éluer les molécules interagissant avec la matrice de la colonne de chromatographie en gradient de NaCl afin de séparer les composés anioniques de faible interaction et de petite taille des molécules de forte avidité pour cette matrice. Deux fractions ont été obtenues par cette méthode (Figure 85A) : une première fraction (12-17mL) élue en faible concentration de sel (0,15-0,6M NaCl pour une conductivité de 21mS/cm à 50mS/cm). Les absorbances aux différentes longueurs d'onde ne saturent pas comparé à une éluion en *step* de sel, ce qui laisse penser à une quantité en contaminants plus faible. L'analyse RMN de cette première fraction d'élution (Figure 85B) montre la présence de composés dont les déplacements chimiques appartiennent aux sucres (la fenêtre spectrale est centrée sur cette région), mais qui ne sont pas des HS. En effet, les déplacements chimiques attendus pour ces derniers, encerclés en vert, ne sont pas observés, indiquant leur absence de l'échantillon ou dans des quantités trop faibles pour être détectées.

La fraction d'élution (18-22mL) avec une concentration en sel plus importante de 0,6 à 0,9M NaCl a une absorbance maximale à 260nm et 280nm (600mAU et 500mAU respectivement) similaire à l'absorbance maximale de la fraction précédente élue en concentration de sel moins importante. L'absorbance à 215nm des liaisons peptidiques est en revanche plus faible (1100mAU) que la fraction précédente (1500mAU) (Figure 85A). L'analyse du spectre RMN de cette fraction d'intérêt montre la présence de contaminants en quantité relativement importante mais également la présence très nette des HS à travers la présence de corrélations spécifiques attendues aux déplacements chimiques des HS (encerclés en vert) (Figure 85C). Les fractions éluées à des concentrations supérieures en sel ne montrent aucune corrélation (non-montré), signifiant que tous les composés sont élués jusqu'à ≈1M NaCl.

L'analyse RMN de la fraction non-retenue (*flowthrough*) ne montre pas la présence d'HS (spectre non montré, similaire à la Figure 84B).

Ensemble, ces deux méthodes de purification montrent que les HS sont retenus par la matrice DEAE avec une forte avidité. En effet, l'analyse RMN des effluents ne montre pas de corrélation attendue pour ceux-ci dans des régions spécifiques du spectre. Les HS sont en fait élués à une concentration saline supérieure à 0,6M NaCl, déterminée grâce à une élution en gradient de sel. L'utilisation d'un gradient de sel permet d'éliminer une partie des contaminants anioniques de faible interaction avec la DEAE ou de petite taille. Les longues chaînes d'HS, mais également d'acides nucléiques nécessitent une concentration en sel plus importante pour casser toutes les interactions électrostatiques. L'utilisation d'un gradient de sel est donc nécessaire afin de disposer d'échantillons d'HS plus purs et plus facilement analysables. Ces échantillons nécessitent cependant une digestion des acides nucléiques à la

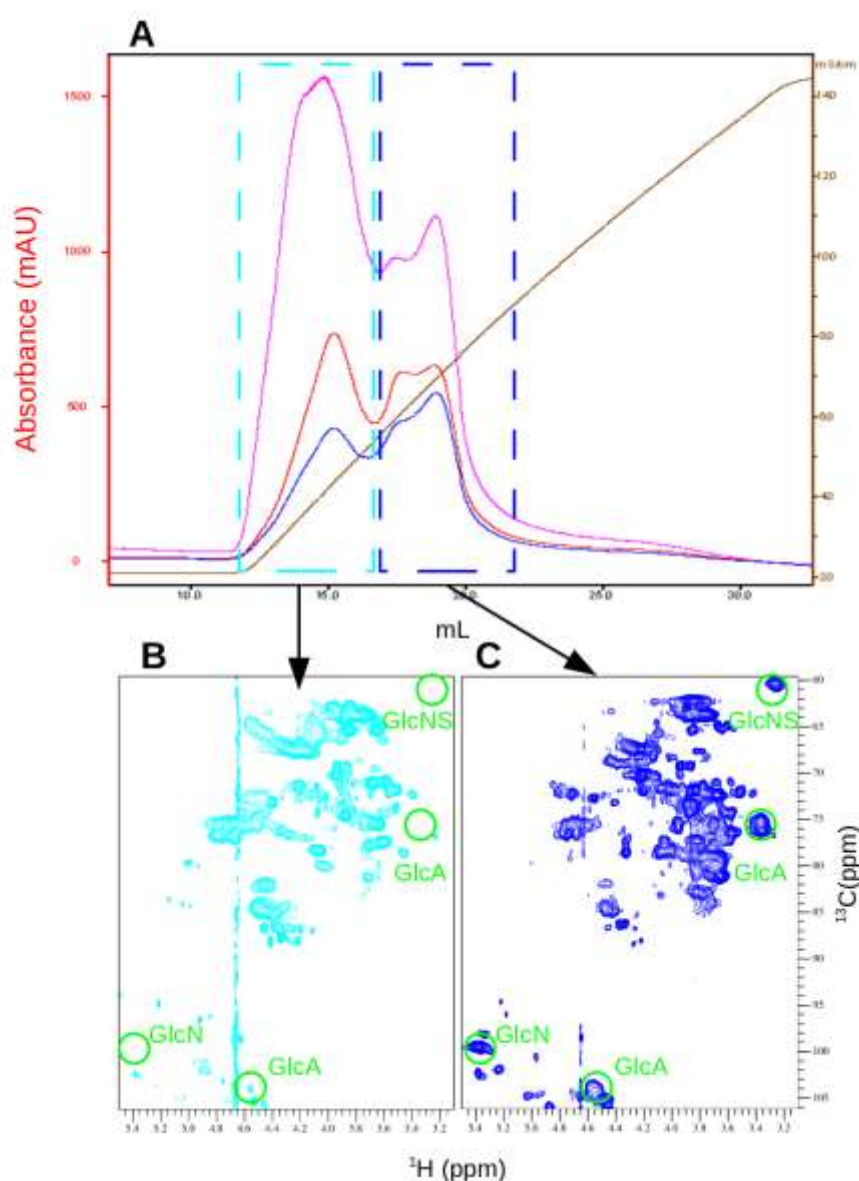


Figure 85 : Purification par gradient de sel et analyse RMN d'une digestion à la trypsine de cellules HeLa. Le chromatographe (A) montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat. Les spectres de RMN ¹H-¹³C HSQC des fractions éluées en faible concentration (B) et en forte concentration NaCl (C) sont partiellement attribués pour les HS

DNase I comme établi précédemment afin de disposer d'échantillons de qualités pour des analyses RMN plus complètes.

Pendant le processus de purification, deux à trois étapes de digestion et de chromatographie sont nécessaires au cours desquelles du matériel d'intérêt peut être perdu. Dans le but de limiter ces pertes d'échantillon, une stratégie de réduction du nombre d'étape a été envisagée.

c) Réduction d'étapes pré-analytiques

Les étapes critiques dans la perte d'échantillon sont principalement les étapes de chromatographie. Afin de ne procéder qu'à une seule chromatographie échangeuse d'anions, les digestions des protéines et des acides nucléiques doivent toutes précéder celle-ci. Suite au traitement à la trypsine des surfaces cellulaires, le digestat produit est lyophilisé puis repris dans 30mL d'eau. La solution est ensuite dialysée. Lors de cette étape, des précipités sont formés et éliminés par centrifugation. Le dialysat est ensuite concentré pour se placer dans de bonnes conditions de digestion enzymatique à la DNase I. A la suite de cette digestion l'échantillon est alors purifié par chromatographie échangeuse d'anions.

Les molécules ayant interagit avec la matrice DEAE, sont éluées en gradient de sel comme déterminé précédemment. Le chromatographe présente deux pics peu résolus (Figure 86A). Dans le premier pic (10-16mL) une absorbance de 100mAU est mesurée à 260nm. Cette absorbance est plus faible que l'absorbance du second pic (17-23mL) qui est de 280mAU.

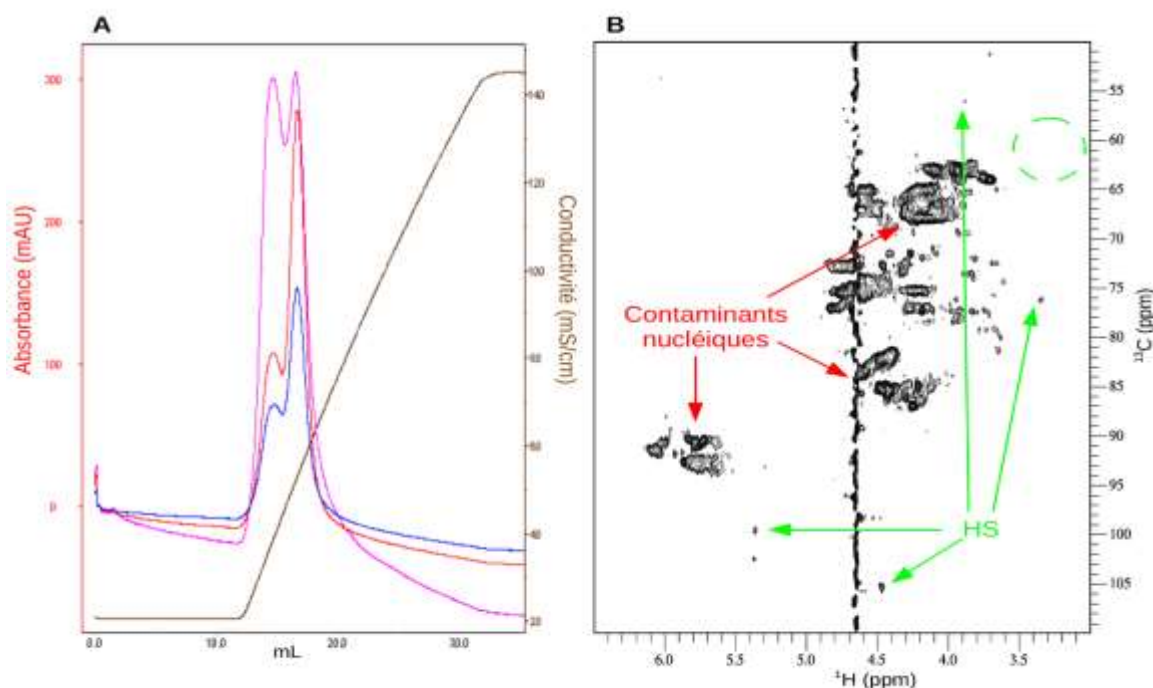


Figure 86 : Purification et analyse RMN d'une digestion à la trypsine puis DNase I de cellules HeLa. Le chromatographe en gradient de sel (A) montre l'absorbance à 260nm (rouge), 280nm (bleu), 215nm (magenta) ainsi que la conductivité (marron) du digestat. Le spectre de RMN ¹H-¹³C HSQC (B) est partiellement attribué et indique les corrélations observées et attendues pour les HS (flèches et cercle verts) et pour les acides nucléiques (flèches rouges).

L'absorbance des protéines/peptides est de 300mAU à 215nm pour chacun des deux pics et de 60mAU et 150mAU à 280nm pour les pics 1 et 2 respectivement.

L'analyse RMN de la fraction de 10 à 16mL ne montre aucune corrélation observable dans nos conditions spectroscopiques (non-montré). La fraction correspondant au second pic (17-23mL) contenant les molécules de forte interaction électrostatique, en revanche, montre des corrélations observables par RMN (Figure 86B). Toutefois, les déplacements chimiques observés correspondent majoritairement à ceux des oses des acides nucléiques indiqués en rouge. Les corrélations attendues pour les HS indiquées en vert ne sont pas observées ou avec une sensibilité très basse.

En résumé, la stratégie mise en place afin de réduire le nombre d'étapes de chromatographie s'est montrée peu efficace pour obtenir des HS purs et en quantité suffisante. Des stades lors de la préparation de l'échantillon semblent être critiques dans la perte d'échantillon. Une de ces étapes est la dialyse où des précipités se forment, pouvant entraîner les molécules d'intérêt. Un second essai réalisé pour confirmer le résultat n'a pas donné de résultats probants.

Le protocole de purification d'HS cellulaires mis en place consiste dans un premier temps en l'extraction d'HS par digestion enzymatique à partir des cellules marquées au ^{13}C . Le nombre de cellules marquées, nécessaires à cette préparation est de 200 millions de cellules, cultivées dans trois flasques T600. Les extraits solubles obtenus après traitement à la trypsine et contenant les ^{13}C -HS sont ensuite purifiés par chromatographie échangeuse d'anions en gradient de sel. Les fractions récoltées dans lesquelles se trouvent les HS sont digérées à la DNase I afin d'éliminer les contaminants des acides nucléiques, les riboses et désoxyriboses, également marqués et visibles par RMN dans la même fenêtre spectrale. Les échantillons sont une nouvelle fois chargés sur colonne échangeuse d'anions et élués en gradient de sel. La fraction d'élution à forte concentration saline présente les corrélations attendues pour les HS avec une réduction nette de la quantité d'ADN/ARN et permet une analyse qualitative et quantitative de la structure des HS.

5. Quantification des HS

Afin de caractériser les échantillons purifiés, il est important de déterminer assez précisément la quantité d'HS récoltés à la surface membranaire des cellules HeLa. Toutefois il n'existe pas de méthodes simples et fiables pour l'analyse des HS dans ces conditions complexes. Cependant, il existe différentes techniques utilisées qui ont été testées et qui seront présentées pour un même extrait issu de digestion à la trypsine des surfaces cellulaires et purifié une première fois par chromatographie échangeuse d'anions.

a) Dosage des acides uroniques par la méthode au carbazole

Le dosage au carbazole (Bitter & Muir 1962) repose sur la fixation de ce composé coloré sur un dérivé des acides uroniques (qui composent les HS, CS et HA) créés dans certaines conditions (voir matériel et méthode). L'absorbance à 530nm est donc

proportionnelle à la quantité en acide uronique. La réactivité croisée des autres GAG semble limitée car seuls les HS sont observés sur les surfaces cellulaires par cytométrie en flux (Figure 68).

Une gamme étalon est réalisée par dilution en cascade d'un échantillon d'HS de concentration connue. Les valeurs d'absorbance à 530nm sont données en fonction de la concentration en HS et une droite d'équation de la forme $y = ax + b$ correspondant à cette gamme est tracée (Figure 87). Le coefficient de détermination R^2 est calculé ; plus sa valeur tend vers 1 et plus la droite est représentative du modèle (ici l'absorbance à 530nm en fonction de la concentration en HS).

L'absorbance de l'échantillon mesurée est de 0,53 ce qui représente une concentration de 100µg/mL d'après la gamme étalon. Sachant que le volume totale est de 400µL, la quantité en HS (et CS potentiels) est de l'ordre de 40µg issus de 200 millions de cellules HeLa. L'analyse a également nécessité 50µL d'échantillon, soit 13% du matériel obtenu (≈5µg).

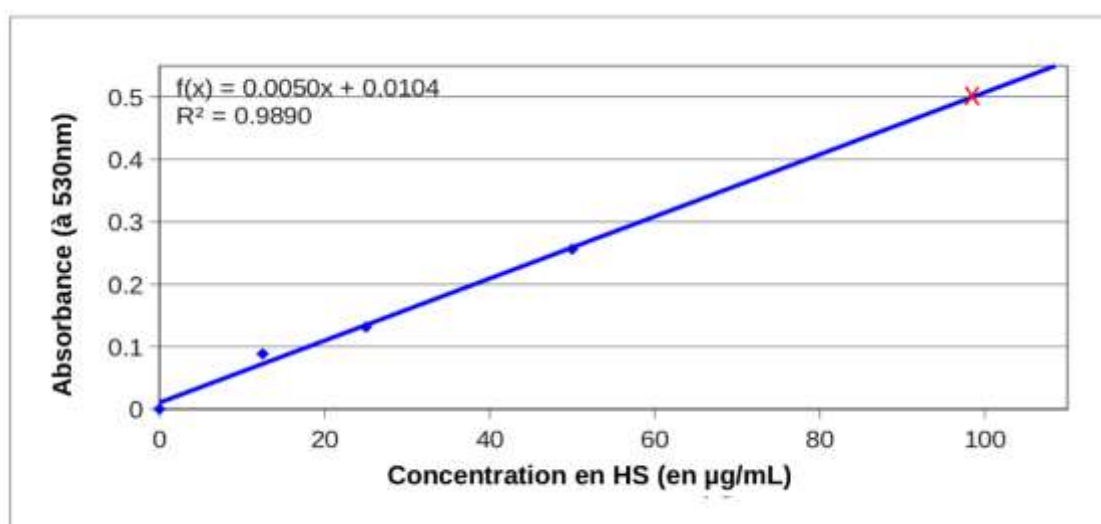


Figure 87 : Dosage par carbazole de l'échantillon pour déterminer la concentration d'HS cellulaires entiers issus de cellules HeLa digérées à la trypsine. La croix rouge correspond à l'absorbance de l'extrait à 530nm reportée sur la droite de régression linéaire.

La concentration en HS de l'extrait a également été estimée par une seconde méthode, commerciale, dans le but de quantifier les HS dans l'extrait et de comparer les deux méthodes.

b) Dosage des GAG sulfatés par la méthode Blyscan™ commercial

Cette technique repose sur l'interaction entre le colorant "Blyscan" et les groupements sulfates des GAG (voir partie Matériel et Méthodes et Figure 67). Tous les sulpho-GAG (sGAG) sont donc quantifiés en fonction de leur quantité en sulfates.

Comme précédemment, une gamme étalon en quantité de sGAG connue est établie en fonction de leur absorbance à 630nm. Une droite est tracée à partir des valeurs obtenues et le coefficient de détermination (R^2) est calculé à partir de ces mêmes valeurs. Le R^2 obtenu (0,9959) (Figure 88) est proche de 1 ce qui signifie que l'absorbance à 630nm est proportionnelle à la quantité de sGAG dans l'échantillon.

L'absorbance à 630nm de l'extrait d'HS issu de la digestion à la trypsine de 200 millions

de cellules HeLa est lue dans les mêmes conditions que pour la gamme étalon et est de 0,249 pour un volume de 5µL ; c'est-à-dire que pour l'extrait de 100µL une estimation de 25µg en sGAG est obtenue. Cette estimation est inférieure à celle donnée par la méthode de dosage au carbazole, ce qui peut s'expliquer en partie par la perte d'échantillon utilisée pour le précédent dosage. Toutefois, cette estimation est d'autant plus proche de la réalité que la quantité de sulfate par chaîne pour les sGAG de la gamme étalon est similaire à celle des HS de l'extrait cellulaire. Les sGAG de la gamme étalon étant en réalité des CS-A (C4S), la quantification en HS de l'extrait est donc surestimée car les HS sont généralement plus sulfatés que les CS.

Le colorant Blyscan se servant des propriétés électro-négatives des composés pour se fixer, la fiabilité de la méthode a été testée en fixant du colorant sur l'ADN qui est une molécule poly-phosphatée. Une quantité de 3µg d'ADN entier et digéré ont été mis en présence du colorant dans les mêmes conditions que celles décrites pour les extraits cellulaires. La quantification de l'ADN non fragmenté est totale alors qu'environ un tiers (1,18µg) du colorant a pu se fixer sur le digestat d'ADN.

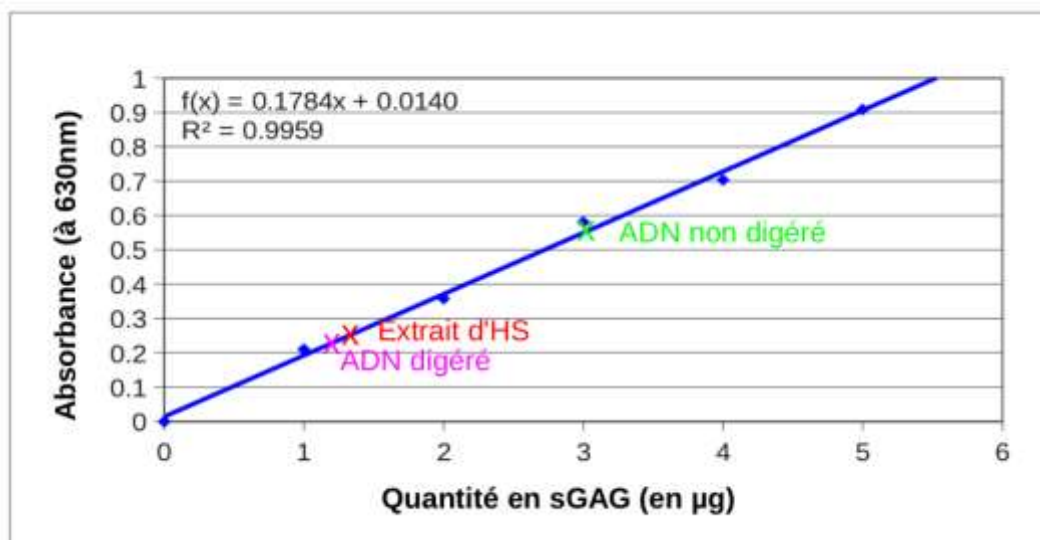


Figure 88 : Dosage par blyscan de l'échantillon pour déterminer la quantité d'HS cellulaires entiers issus de cellules HeLa digérées à la trypsine. Sur la droite de régression linéaire sont reportées les absorbances à 630nm de l'extrait d'HS cellulaires (croix rouge), d'ADN non-digéré (croix verte) et d'ADN digéré (croix magenta).

Ensemble ces résultats suggèrent que le dosage au colorant Blyscan est peu spécifique et nécessite un échantillon pur et une connaissance préalable de la sulfatation de l'extrait afin d'établir une gamme étalon adéquate. Nos extraits de cellules comprenant toujours une certaine quantité en acides nucléiques, il n'est pas possible de calculer directement la quantité en HS dans nos digestats cellulaires.

Enfin, une dernière méthode de quantification, par spectroscopie RMN, a également été mise en place pour estimer la quantité en HS cellulaires.

c) Dosage des HS par RMN

La détermination de la quantité en HS des cellules par RMN est réalisée par comparaison avec un échantillon d'HS de masse connue. Comme vu dans la partie

introduction, l'intensité d'un signal en RMN dépend du temps expérimental (qui lui-même dépend du nombre d'itérations de l'expérience RMN dans les différentes dimensions considérées), du gain et de la quantité en noyaux observables.

Un spectre est enregistré à partir d'un échantillon d'HS commercial de 2mg non-marqué. La quantification est réalisée par l'intégration du volume des pics de corrélation ^1H - ^{13}C des glucosamines. Le signal observé dépend donc de la probabilité d'avoir un ^1H (99,99%) directement couplé avec un ^{13}C (1,1%) (Figure 54) soit une probabilité de 1,1%. Sur les 2mg d'HS commerciaux, seulement 20 μg sont donc observables. Le volume des signaux des glucosamines observées est 8,2 pour un temps expérimental de 13h15mn (Figure 89).

Notre extrait issu de 200 millions de cellules HeLa, en considérant qu'il est marqué à 100% en ^{13}C , analysé en 16h à un volume de pic au niveau des glucosamines de 3,8. En prenant en compte ce volume en fonction du temps d'expérience comparé à l'échantillon d'HS connu, une dizaine de microgrammes d'HS est ainsi estimée (Figure 89). La quantification par RMN offre l'avantage de pouvoir intégrer des pics appartenant spécifiquement aux HS, et ainsi de ne pas quantifier des contaminants, comme cela est le cas avec le dosage au blyscan. Des biais peuvent cependant être introduits du fait des conditions spectroscopiques différentes (temps expérimental, champ magnétique...) et expérimentalement par les différentes limites d'intégration des pics observés.

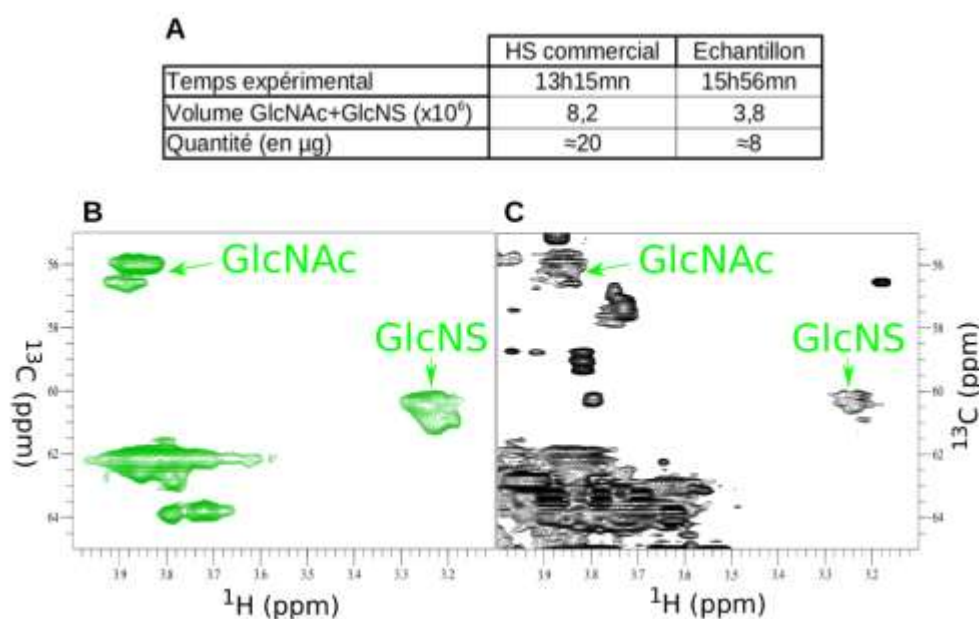


Figure 89 : Dosage par RMN d'une fraction d'HS issus de cellules HeLa comparativement à un échantillon commercial de masse connue. Les temps expérimentaux, volumes des pics de corrélation $\text{H}_2\text{-C}_2$ des glucosamines et des quantités en HS connus ou calculés sont donnés en A. Les spectres HSQC au niveau des pics intégrés sont montrés pour les HS commerciaux (B) et pour l'échantillon de masse inconnu (C).

En résumé, la quantification des HS cellulaires n'est pas aisée du fait de la complexité de l'échantillon et des réactions ou conditions parasites qui peuvent biaiser les résultats obtenus. Cependant, les différentes méthodes de dosage colorimétrique et spectroscopique permettent d'estimer à environ une dizaine de microgrammes la quantité d'HS obtenue après digestion, extraction et purification à partir de 200 millions de cellules HeLa. Les analyses RMN

montrent également une quantité non-négligeable d'acides nucléiques dont leur élimination est nécessaire pour pouvoir mettre en place des analyses plus qualitative et quantitative par RMN.

B. Analyses par RMN en solution des HS purifiés issus des cellules HeLa

Dans la partie précédente, j'ai montré qu'il était possible de purifier des quantités suffisantes d'HS cellulaires pour enregistrer des spectres ^{13}C -HSQC dans un temps raisonnable, le protocole d'extraction ayant été optimisé. Pour leur étude par RMN, différents paramètres peuvent influencer la précision de l'analyse. La quantité de matériel tout d'abord, qui est indépendante de l'expérimentateur, mais la température d'analyse est un paramètre ajustable qui a un effet sur l'agitation moléculaire, donc sur le τ_c et donc sur le temps de relaxation transversale.

1. Influence de la température d'analyse

En premier lieu, l'influence de la température sur la qualité spectrale a été étudiée. A priori, les HS sont des molécules flexibles dont les signaux sont déjà fins. Toutefois, afin de déterminer si la température a un effet significatif sur la largeur des signaux et donc sur leur intensité, trois températures d'analyse de 25°C, 37°C et 50°C sont testées. Sur la Figure 90, des extraits des spectres HSQC de la zone des noyaux anomériques des HS sont présentés pour chaque température d'analyse dans les mêmes conditions spectroscopiques pour le même échantillon.

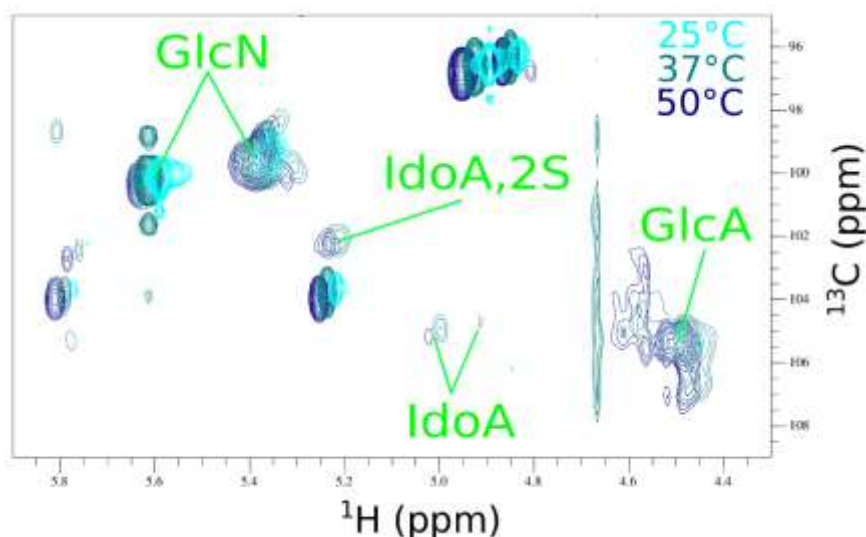


Figure 90 : Analyse par RMN d'un échantillon d'HS cellulaires à différentes températures dans la zone de corrélation des noyaux anomériques. Les HS extraits et purifiés sont analysés avec des expériences de corrélation ^1H - ^{13}C HSQC à 25°C (cyan), 37°C (bleu canard) et 50°C (bleu marine) dans un spectromètre 600MHz équipé d'une cryosonde pendant 16h.

Le spectre enregistré à 25°C (en cyan) montre clairement la présence des corrélations des glucosamines (GlcN) à 5,35-99,6ppm et 5,6-100ppm ^1H - ^{13}C . Cependant les signaux attendus pour les acides uroniques sont de faible intensité pour les GlcA (4,47-105ppm ^1H -

^{13}C) et au niveau du bruit pour les acides iduroniques. L'analyse de cet échantillon d'HS dans ces conditions spectroscopiques permet de rendre compte de la présence effective des HS. Toutefois, le rapport signal/bruit de l'expérience ne permet pas ici de quantifier correctement les différentes espèces observées.

La température au sein de l'échantillon est augmentée jusqu'à 37°C (pics bleu canard). Tous les signaux observés à 25°C sont également observés à 37°C. Comme cela peut être attendu de l'affinement des signaux, l'intensité des pics est améliorée. C'est le cas au niveau des glucosamines (5,35-99,6ppm ^1H - ^{13}C) mais de manière plus significative au niveau des GlcA (4,47-105ppm ^1H - ^{13}C) dont le rapport signal/bruit est multiplié par 7. La sensibilité de l'expérience dans ces conditions de température permettent dorénavant d'observer et de distinguer les acides iduroniques (5,22-102,2ppm ; 5-105ppm ; 4,9-104,8ppm ^1H - ^{13}C) du bruit de fond.

Nous avons ensuite augmenté jusqu'à 50°C la température à laquelle le spectre de l'échantillon d'HS est enregistré (pics bleu marine). Tous les pics observés à 37°C sont également observés à 50°C. Par contre, cette augmentation de température ne permet pas d'observer des pics d'HS supplémentaires par rapport au spectre HSQC enregistré à 37°C. De même, les intensités des signaux des spectres enregistrés à 37°C et 50°C sont similaires.

Cette analyse nous montre l'influence significative que joue la température sur la qualité des spectres obtenus par RMN, en permettant d'observer des pics de faible intensité à partir de 37°C alors qu'ils restaient indétectables à 25°C dans des conditions spectroscopiques identiques par ailleurs. Une augmentation supplémentaire de la température jusqu'à 50°C ne montre pas en revanche une augmentation significative du rapport signal/bruit. Dans ces conditions, nous nous sommes fixés une température de 37°C pour toutes les expériences qui ont été enregistrées dans la suite. Enregistrer les données à plus haute température ne permettrait pas un gain du signal tout en laissant la possibilité d'une dégradation des HS au cours du temps.

2. Attribution des signaux

Le spectre ^{13}C -HSQC enregistré à 37°C d'un échantillon représentatif d'HS purifiés à partir de cellules HeLa est présenté dans la figure Figure 91. Une première observation est que le spectre présente tous les signaux attendus pour des chaînes d'HS. Il peut être également observé des signaux supplémentaires qui correspondent à la présence de chaînes de CS mais de faible intensité. Enfin quelques signaux ne correspondant ni aux HS, ni aux CS sont également présents mais sont aussi de faible intensité.

Les attributions des signaux des HS sont réalisées par comparaison avec les données de la littérature données principalement par Guerrini et al 2005.

Deux régions d'intérêt peuvent être distinguées où les différentes modifications des oses peuvent être caractérisées. La première est la région correspondant aux corrélations des noyaux ^1H et ^{13}C anomériques ($^1\text{H}_1$ - $^{13}\text{C}_1$) (4-6ppm en ^1H et 95-110ppm en ^{13}C , spectre en bas à droite). De ce fait c'est une région ne présentant que peu de recouvrements spectraux et où l'influence d'une (ou plusieurs) modification de l'ose, ou de celle(s) de l'ose voisin peut être

observée sur le déplacement chimique de la paire de noyaux $^1\text{H}_1\text{-}^{13}\text{C}_1$. C'est le cas par exemple pour les $^1\text{H}_1\text{-}^{13}\text{C}_1$ des GlcA. Lorsque celui-ci est suivi d'une glucosamine non-modifiée (GlcNAc), le déplacement chimique en ^1H est de 4,49ppm alors qu'il est de 4,59ppm lorsqu'il est suivi d'une GlcNS pour un déplacement chimique comparable en ^{13}C (105ppm). Quand le GlcA est épimérisé en IdoA (requérant une GlcNS adjacente), les déplacements chimiques sont plus déblindés (déplacements vers les bas champs), au-delà de 4,9ppm en proton. La présence des groupements 6-O-sulfates de la GlcNS adjacente influence également le déplacement chimique en proton des IdoA (4,93ppm quand IdoA relié à une GlcNS,6OH et 5,00ppm quand relié à une GlcNS,6S). L'épimérisation influence peu le déplacement chimique du carbone ($\approx 104,7\text{ppm}$). En revanche, la 2-O-sulfatation de l'IdoA provoque un changement d'environnement chimique traduit par des déplacements chimiques différents en proton (5,25ppm) et en carbone (102ppm). Dans ce cas, la modification de l'ose directement lié n'a

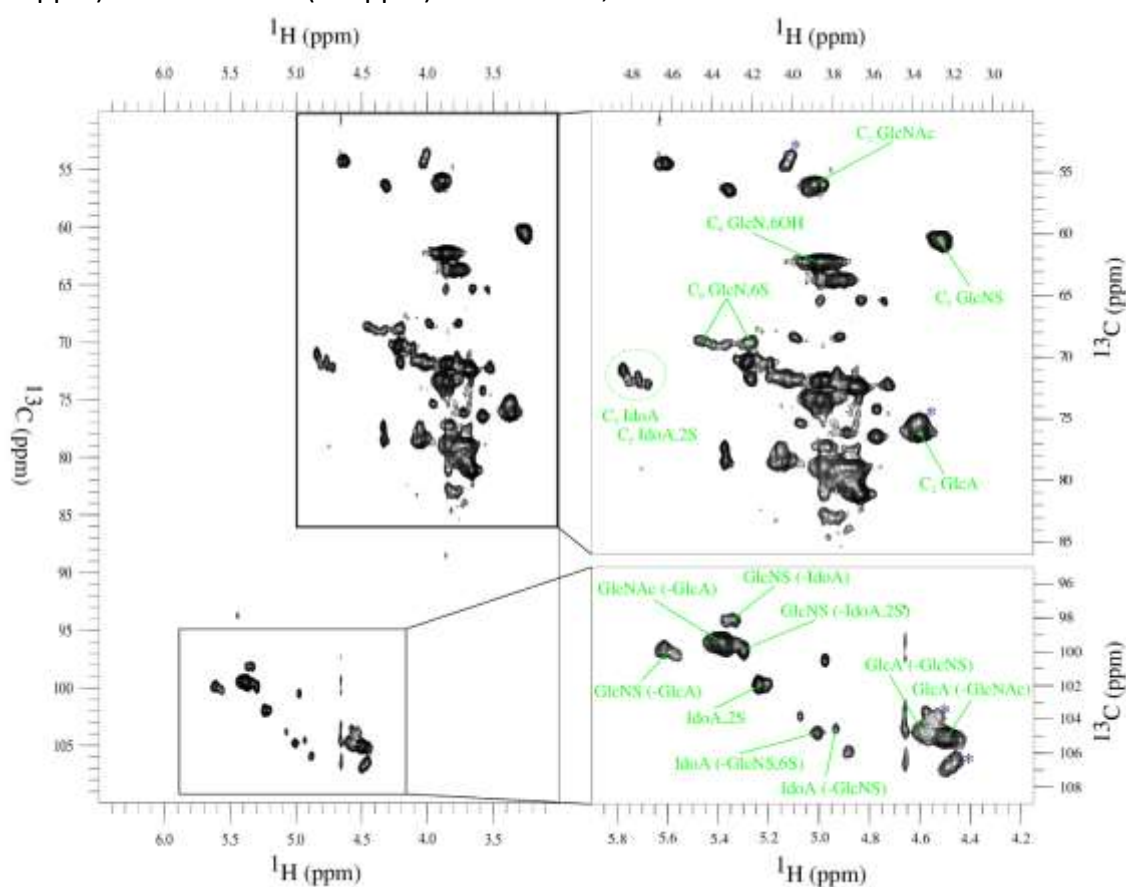


Figure 91 : Attribution d'un spectre RMN ^{13}C -HSQC d'HS cellulaires purifiés selon le protocole précédemment décrit. Attribution des HS indiqués en vert d'après Guerrini et al 2005. La présence des corrélations appartenant aux CS est indiquée (étoiles bleues) d'après Huckerby et al 2005. L'échantillon est analysé à 37°C pendant 16h dans un spectromètre 600MHz équipé d'une cryosonde.

pas d'influence sur le déplacement chimique.

Concernant les corrélations $^1\text{H}_1\text{-}^{13}\text{C}_1$ des glucosamines, lorsque celles-ci ne sont pas modifiées, les déplacements chimiques sont de 5,38ppm en ^1H et de 99,5ppm en ^{13}C . Lors de la N-sulfatation, les déplacements chimiques sont différents de ceux des glucosamines non-sulfatées mais vont dépendre essentiellement de la (ou des) modification(s) du sucre voisin (épimérisation avec/sans 2-O-sulfatation, pas de modification). Par contre, la 6-O-sulfatation

des glucosamines n'induit pas d'effet notable sur le déplacement chimique des corrélations $^1\text{H}_1\text{-}^{13}\text{C}_1$ et donc ces espèces ne peuvent pas être caractérisées dans cette région du spectre.

Enfin, une seconde région du spectre de RMN regroupant les autres noyaux du cycle osidique peut être observée ($^1\text{H}_{2-6}$ et $^{13}\text{C}_{2-6}$). Cette région présente un grand nombre de superpositions, les atomes des oses ayant un déplacement chimique similaire du fait leur environnement chimique proche (par exemple du fait que les carbones soient tous reliés à un atome d'oxygène et un atome de carbone) et seules quelques corrélations spécifiques y sont suffisamment résolues.

Les corrélations $^1\text{H}_2\text{-}^{13}\text{C}_2$ des GlcA sont observées à 3,37-76ppm ($^1\text{H}\text{-}^{13}\text{C}$) mais peu résolues entre ceux des CS et des HS, ces derniers étant un peu plus déblindés dans la dimension ^{13}C . Quand les GlcA sont épimérisés en IdoA, seules les corrélations $^1\text{H}_5\text{-}^{13}\text{C}_5$ des IdoA et IdoA2S sont observées dans la zone 4,8-71,5ppm ($^1\text{H}\text{-}^{13}\text{C}$).

Concernant les glucosamines, les GlcNAc et les GlcNS présentent des corrélations $^1\text{H}_2\text{-}^{13}\text{C}_2$ dans une zone du spectre peu encombrée. Par contre ces corrélations sont très peu sensibles à la nature de l'acide uronique adjacent et forment deux massifs non-résolus à 3,9-56ppm et 3,25-60,5ppm ($^1\text{H}\text{-}^{13}\text{C}$) respectivement. Les formes 6-*O*-sulfatées et non-modifiées sont également observées à 3,85-62,5ppm ($^1\text{H}\text{-}^{13}\text{C}$) pour les glucosamines 6-hydroxylées et à 4,35-69ppm ($^1\text{H}\text{-}^{13}\text{C}$) pour celles 6-*O*-sulfatées.

Ainsi, nous avons pu voir que l'échantillon purifié d'HS cellulaires pouvait être analysé de manière qualitative par RMN bidimensionnelle. En effet, les différentes corrélations attendues sont présentes et leur intensité et résolution sont suffisantes pour pouvoir faire des analyses quantitatives dans ces conditions spectroscopiques.

3. Quantification des signaux

La structure des HS est liée à ses différentes modifications des GlcA et GlcNAc qui les composent. A partir des intensités relatives (intensité d'un pic par rapport à un autre pic) des différents oses modifiés, il est alors possible d'en déterminer son taux de modification et son profil disaccharidique.

a) La N-sulfatation

Différentes méthodes peuvent être utilisées pour déterminer le taux de N-sulfatation des HS de l'échantillon. Une première technique serait d'intégrer toutes les intensités des différentes corrélations des GlcNS : GlcNS-GlcA, GlcNS-IdoA et GlcNS-IdoA,2S et des GlcNAc de la région des noyaux anomériques et d'en faire le rapport $\frac{\text{GlcNS}}{\text{GlcNAc} + \text{GlcNS}} \times 100$. Toutefois, du fait de l'éclatement des pics, cela augmente l'incertitude sur l'intensité. D'autre part, le pic $^1\text{H}_1\text{-}^{13}\text{C}_1$ des glucosamines de l'enchaînement GlcNS,6S-IdoA,2S (Yates et al. 1996) est confondu avec celui des GlcNAc ce qui conduirait à sous-estimer le taux de sulfatation.

C'est pour cela que le taux de N-sulfatation est déterminé à partir des corrélations $^1\text{H}_2\text{-}^{13}\text{C}_2$ (Figure 92). Ces corrélations des glucosamines permettent de limiter le nombre de pics à intégrer et donc de réduire l'imprécision de la mesure. De plus, toutes les formes des GlcNAc

et des GlcNS indépendamment de l'ose adjacent ont un même déplacement chimique ce qui permet d'augmenter le signal au niveau de ces pics, et donc de réduire l'importance du bruit dans la mesure.

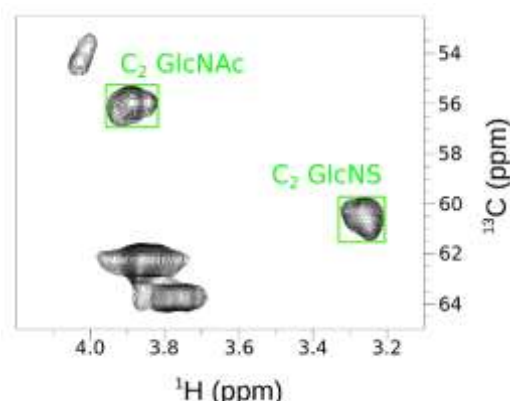


Figure 92 : Spectre RMN ^{13}C -HSQC des déplacements chimiques $\text{H}_2\text{-C}_2$ des glucosamines d'HS. Les encadrés verts représentent les bornes utilisées pour l'intégration des pics. L'échantillon est analysé à 37°C pendant 16h dans un spectromètre 600MHz équipé d'une cryosonde.

Des intégrations sur plusieurs cultures indépendantes (n=3) de cellules HeLa ont été réalisées. Le taux de N-sulfatation calculé est de $46 \pm 2,6\%$. Ce taux de sulfatation est approximativement le même entre les différentes cultures, ce qui montre également la reproductibilité du protocole utilisé ainsi que la faible variation de ce phénotype au cours des différentes cultures (Figure 93).

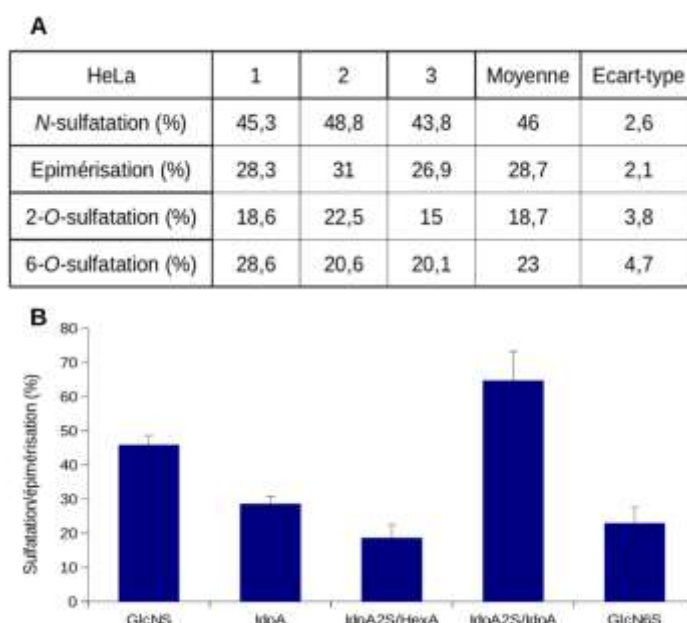


Figure 93 : Caractéristiques des HS de cellules HeLa. Tableau de valeurs (A) et graphique des sulfatations/épimérisation (B) des HS issus de trois cultures cellulaire HeLa indépendantes.

b) L'épimérisation des acides glucuroniques

Le taux d'épimérisation des HS est déterminé par l'intégration de tous les pics des acides iduroniques sur l'ensemble des acides uroniques (GlcA+IdoA notés HexA pour acide hexuronique) $\frac{\text{IdoA2S} + \text{IdoA}(-\text{GlcNS}) + \text{IdoA}(-\text{GlcNS},6\text{S})}{\text{IdoA2S} + \text{IdoA}(-\text{GlcNS}) + \text{IdoA}(-\text{GlcNS},6\text{S}) + \text{GlcA}(-\text{GlcNAc}) + \text{GlcA}(-\text{GlcNS})} \times 100$ dans la

zone de corrélation des noyaux anomériques. Pour les 3 cultures cellulaires HeLa, le taux de C₅-épipimérisation mesuré est de 28,7±2,1%. Comme pour la N-sulfatation, la quantité en acides iduroniques semble stable selon les différentes cultures (Figure 93A).

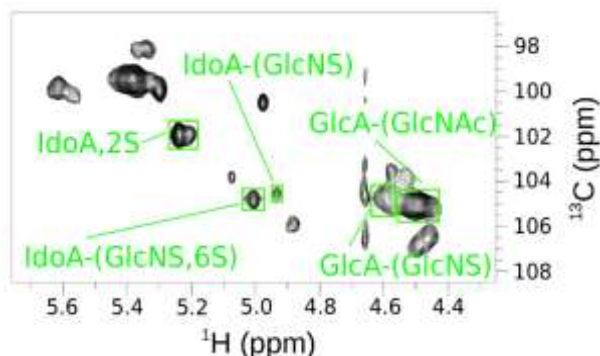


Figure 94 : Spectre RMN ¹³C-HSQC des déplacements chimiques H₁-C₁ des acides uroniques d'HS. Les encadrés verts représentent les bornes utilisées pour l'intégration des pics. L'échantillon est analysé à 37°C pendant 16h dans un spectromètre 600MHz équipé d'une crysonde.

c) La 2-O-sulfatation des acides iduroniques

La 2-O-sulfatation est déterminée à partir du signal des IdoA,2S comparé aux autres acides iduroniques ou uroniques $\frac{IdoA2S}{\sum IdoA} \times 100$ ou $\frac{IdoA2S}{\sum HexA} \times 100$ dans la zone des noyaux anomériques (Figure 94). Au sein d'une chaîne d'HS, en moyenne 18,7±3,8% des acides uroniques sont 2-O-sulfatés, et sont très majoritairement portés par les IdoA. Les IdoA,2S représentent 64,8±8,6% de tous les IdoA (Figure 93).

d) La 6-O-sulfatation des glucosamines

Le taux de 6-O-sulfatation est déterminé par l'intégration des pics des corrélations ¹H₆-¹³C₆ des glucosamines 6S et 6OH $\frac{GlcN,6S}{GlcN,6S+GlcN,6OH} \times 100$ (Figure 95). La pureté de l'échantillon est importante afin de pouvoir déterminer précisément le taux moyen de 6-O-sulfatation car des corrélations issues des contaminants peuvent parasiter cette région du spectre et rendre

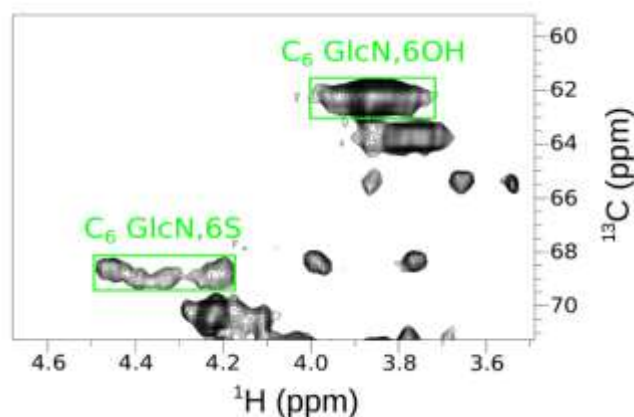


Figure 95 : Spectre RMN ¹³C-HSQC des déplacements chimiques H₆-C₆ des glucosamines d'HS. Les encadrés verts représentent les bornes utilisées pour l'intégration des pics. L'échantillon est analysé à 37°C pendant 16h dans un spectromètre 600MHz équipé d'une crysonde.

impossible l'analyse quantitative. Les chaînes d'HS issues de cellules HeLa présentent un taux de 6-*O*-sulfatation moyen de $23 \pm 4,7\%$ ($n=3$) (Figure 93).

e) La longueur des chaînes

Une estimation de la longueur des chaînes peut être obtenue en utilisant une des corrélations $^1\text{H}_5\text{-}^{13}\text{C}_5$ de l'unité xylose du tétrasaccharide *linker* comme référence interne. En effet, le mode de digestion et de purification des chaînes d'HS que nous avons utilisé doit conserver le tétrasaccharide *linker* dans la chaîne. Donc, chaque chaîne d'HS, en dehors de toute dégradation fortuite, est censée contenir une unité xylosique. L'une des corrélations $^1\text{H}_5\text{-}^{13}\text{C}_5$ présente des déplacements chimiques à (3,4 ppm ^1H - 65,7 ppm ^{13}C) dans une zone résolue du spectre (Figure 96). La comparaison des intensités de cette corrélation – comptant pour un proton – avec celle des signaux $^1\text{H}_2\text{-}^{13}\text{C}_2$ des glucosamines comptant pour autant de protons que d'unités disaccharidiques permet alors de pouvoir estimer la longueur des chaînes selon la formule suivante : $\frac{\text{Xyl}}{\text{GlcNS}+\text{GlcNAc}}$. Ceci ne peut toutefois donner qu'une estimation grossière de la longueur de chaîne dans la mesure où l'on compare un signal a priori faible par rapport à un autre grand. De fait, dans le cas des cellules HeLa analysées ici, les signaux $^1\text{H}_5\text{-}^{13}\text{C}_5$ des xyloses (3,4-65,7ppm) sont peu intenses, proches du bruit de fond. L'estimation que nous pouvons faire ici est très imprécise, de l'ordre de 70 à 100 disaccharides, ce qui correspondrait à des masses moléculaires de 25 à 40 kDa en prenant en compte la contribution des chaînes de CS dans ce calcul comportant également une unité de xylose par chaîne (quantité en CS déterminée à moins de 15% par RMN).

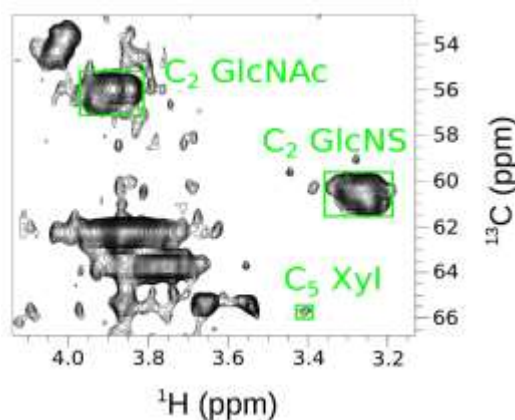


Figure 96 : Spectre RMN ^{13}C -HSQC des déplacements chimiques $\text{H}_2\text{-C}_2$ des glucosamines et d'une paire $\text{H}_5\text{-C}_5$ du xylose d'HS. Les encadrés verts représentent les bornes utilisées pour l'intégration des pics. L'échantillon est analysé à 37°C pendant 16h dans un spectromètre 600MHz équipé d'une crysonde.

f) 2-*O*-sulfatation des GlcA, 3-*O*-sulfatation des GlcN et glucosamines non-substituées

Les acides glucuroniques 2-*O*-sulfatés présentent un signal caractéristique pour les corrélations $\text{H}_2\text{-C}_2$ à 4,14ppm-83,1ppm (Guerrini et al. 2005). De même, les glucosamines 3-*O*-sulfatées présentent un signal caractéristique de la 3-*O*-sulfatation à 5,19ppm-100ppm pour les $\text{H}_1\text{-C}_1$ de 5,19ppm-100ppm (Guerrini et al. 2005). Ces signaux ne sont pas observés dans

les spectres des HS issus de cellules HeLa ce qui suggère que ces sulfatations sont rares dans ce matériel. De même, les corrélations des glucosamines non-substituées (GlcNH_3^+) dont les corrélations $\text{H}_1\text{-C}_1$ sont attendues à 5,43ppm-57ppm ne sont pas observées.

g) Analyse des unités disaccharidiques

De cette analyse, on peut maintenant estimer le nombre moyen de sulfate par disaccharide. Ce nombre est de 0,88 et est la somme de la *N*-sulfatation, de la 2-*O*-sulfatation et de la 6-*O*-sulfatation. D'autre part, il est également possible d'après ces spectres d'obtenir des informations sur les enchaînements disaccharidiques des HS. Les GlcNS sont majoritairement suivies d'un IdoA±2S au moins à 70%. Cette mesure est probablement par défaut, car, comme précisé précédemment, la mesure de la quantité de GlcNS,6S-IdoA,2S n'est pas possible. Les 30% restants (GlcNS-GlcA) correspondraient aux domaines de transition NA-NS. Si l'on estime maintenant à partir de l'intensité des corrélations des GlcA, le pourcentage de glucosamines liées *N*-sulfatées, on obtient de l'ordre de 20%. Cette différence de valeur peut provenir de l'intégration du pic GlcA-GlcNS dans la zone des noyaux anomériques. En effet, les pics des GlcA-GlcNS et GlcA-GlcNAc étant mal résolus, le choix des limites d'intégration reste arbitraire.

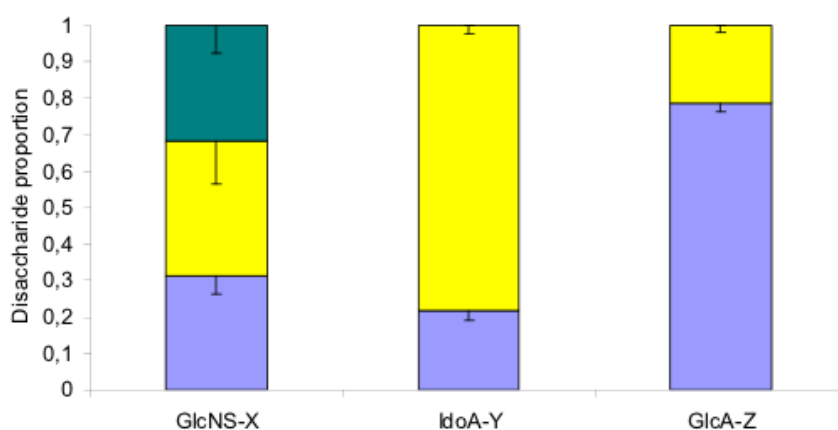


Figure 97 : Caractéristiques des unités disaccharidiques des HS de cellules HeLa. Proportion des disaccharides des HS de cellules HeLa dans lesquels une $\text{GlcNS}\pm 6\text{S}$ est associée à X = GlcA (bleu), IdoA (jaune) ou IdoA,2S (vert) ; un IdoA est associé à Y = GlcNS (bleu) ou GlcNS,6S (jaune) ; un GlcA est lié à Z = GlcNAc (bleu) ou GlcNS (jaune).

Les GlcNS sont en revanche suivies de manière égale par les IdoA ($\approx 35\%$) et IdoA,2S ($\approx 35\%$). Toutefois, les IdoA sont majoritairement suivis de GlcNS,6S (80%) comparées aux GlcNS (20%) (Figure 97).

Ces résultats sont en accord avec les données existantes sur les HS de Gallagher et al. 1985 estimant la *N*-sulfatation des HS de mammifères de différents tissus entre 40 et 50 %. Pour les cellules HeLa, nous avons quantifié le nombre de *N*-sulfates à 46 % en moyenne, là aussi en accord avec Takegawa et al. 2011 pour ce type cellulaire. Le ratio *N*-/*O*-sulfates est généralement supérieur à 1 (Gallagher & Walker 1985), et pour les cellules HeLa ici étudiées ce ratio est de 1,1.

Ainsi nous avons montré que les chaînes d'HS extraites des protéoglycanes et des surfaces cellulaires pouvaient être purifiées et analysées par des techniques de RMN liquide. En effet, les signaux obtenus présentent une sensibilité et une résolution suffisante pour

détecter et discriminer une grande quantité d'unités disaccharidiques diversement modifiées. Toutefois, les modifications rares telles que les 2-*O*-sulfatations des GlcA et les 3-*O*-sulfatations ou non-substitutions des GlcN ne sont pas présentes dans ces HS en quantité suffisante pour être observables. La sensibilité de l'expérience permet cependant d'estimer la longueur moyenne des chaînes d'HS des cellules HeLa présentes à leur surface grâce à la détection d'une corrélation du xylose qui n'est pourtant présente qu'en un seul exemplaire par chaîne.

C. Etudes de polysaccharides d'HS issus de deux types cellulaires : application du protocole et analyse par RMN liquide

Dans les parties précédentes, j'ai développé un protocole de marquage et de purification des HS qui permet d'obtenir à partir de 200 millions de cellules HeLa une dizaine de microgrammes d'HS marqués au ^{13}C . Cette quantité est suffisante pour l'acquisition en un temps raisonnable (de 12 à 16 heures sur un spectromètre 600MHz) d'un spectre ^{13}C -HSQC d'une qualité permettant une analyse des modifications (sulfatations et épimérisation) de ces chaînes. Dans cette partie, je vais présenter l'application de ce protocole à deux autres modèles cellulaires. Pour valider ce protocole, je vais comparer les HS extraits des cellules MEF (fibroblastes de souris embryonnaires) sauvages et des cellules MEF où l'une des enzymes de la voie de synthèse des HS est supprimée. D'autre part, grâce à ce protocole, je vais également analyser les variations des modifications des HS apportées au cours de la différenciation cellulaire sur des cellules épithéliales de type Caco-2.

1. Les fibroblastes de souris embryonnaires (MEF)

Les MEF, et plus généralement les fibroblastes, sont les cellules productrices majeures de la matrice extracellulaire en sécrétant en abondance notamment les PG et les collagènes. Ces fibroblastes sont donc d'une grande importance pour l'organisme formé ou en cours de formation en étant responsables de l'établissement de structures solides et solubles pour les différentes cellules et leur développement.

Ici, je vais comparer les HS issus de la lignée sauvage aux HS issus de la lignée MEF *KO* pour le gène codant une enzyme clé de la biosynthèse des HS : la NDST-1 (cellules notées *ndst-1^{-/-}*). L'enzyme est responsable de la *N*-désacétylation/*N*-sulfatation des GlcNAc en GlcNS et forme les substrats préférentiels pour les enzymes plus en aval dans la voie de biosynthèse des HS. Nous avons donc voulu caractériser l'impact de la suppression de la NDST-1 sur la biosynthèse des HS et les différences structurales des chaînes issues de ces cellules.

a) Détermination par cytométrie en flux de la présence d'HS au niveau des surfaces cellulaires

Les cellules sauvages (MEF WT) incubées avec l'anticorps dirigé contre les HS montrent une intensité de fluorescence importante (médiane de 100, courbe rouge) comparée au contrôle isotypique (médiane de 4, courbe bleue) ce qui indique la présence d'HS en quantité

importante comme pour cellules HeLa (Figure 98A). La présence d'HS au niveau des surfaces cellulaires de MEF *ndst-1*^{-/-} est également observée et nous pouvons constater une diminution dans l'intensité de fluorescence (médiane à 60) (Figure 98B). Ce résultat ne signifie pourtant pas que la cellule *ndst-1*^{-/-} produit moins de chaînes d'HS mais plutôt qu'il y a effectivement un changement de structure au niveau des HS lorsque les cellules n'expriment plus la NDST-1.

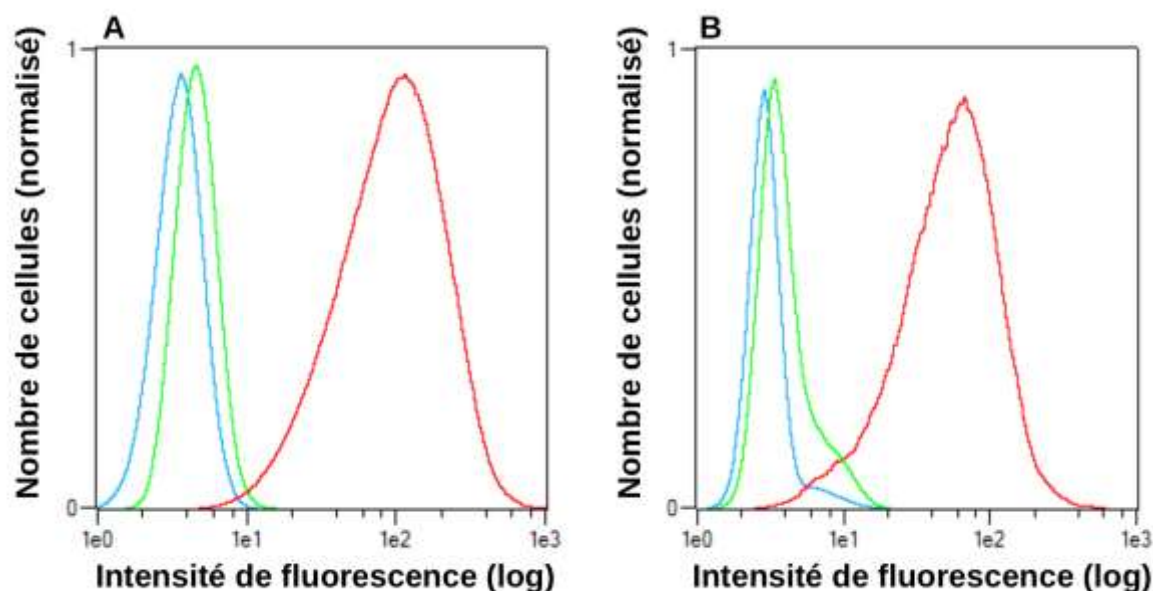


Figure 98 : Analyse par cytométrie en flux des cellules MEF WT (A) et *ndst-1*^{-/-} (B). Des anticorps spécifiques sont utilisés pour détecter les GAG à la surface des cellules. Détection des HS (rouge) et CS (vert), contrôle isotypique (bleu).

Des techniques plus qualitatives doivent être mises en place pour caractériser ces changements.

Les cellules WT et KO marquées aux anticorps dirigés contre les CS (courbes vertes Figure 98) montrent une intensité de fluorescence légèrement supérieure à celle du contrôle isotypique. Cela suggère une présence des CS à la surfaces de ces deux variants cellulaires (WT et KO) dans des quantités très faibles par rapport aux HS. Notons néanmoins que nous avons observé des intensités de fluorescence similaires chez les cellules HeLa pour une présence de CS de l'ordre de 10 % d'après l'analyse des résultats de RMN. Cela suggère que l'absence de la NDST-1 ne joue pas de rôle au niveau de l'expression des chaînes des CS.

Ainsi, les résultats obtenus par cytométrie en flux montrent la présence d'HS en quantité moindre sur les surfaces des MEF WT et KO par rapport aux cellules HeLa. Toutefois, Les MEF, par la présence encore importante en HS, constituent un bon modèle d'étude, et la caractérisation des chaînes d'HS peut être menée pour évaluer l'impact de l'absence de la NDST-1 sur la structure finale de la chaîne d'HS.

b) Marquage et purification

Les cellules MEF sont cultivées en milieu marqué contenant le ¹³C-glucose comme pour les cellules HeLa. Les HS sont issus de 200 millions de cellules après digestion à la trypsine à 37°C.

c) *Analyse RMN de l'évolution du profil saccharidique*

Dans la Figure 99 sont représentés les spectres ^{13}C -HSQC des fractions d'HS issus des cellules MEF sauvages et des cellules MEF *ndst-1^{-/-}*. Les spectres sont globalement équivalents sur le plan de la qualité spectrale. En particulier, ils ne présentent pas de contamination notable par des acides nucléiques et permettent tous les deux une analyse quantitative d'un certain nombre de caractéristiques structurales des HS. Un premier examen montre pour les cellules MEF *ndst-1^{-/-}* la diminution de certains des pics associés aux espèces sulfatées.

Dans la Figure 99 sont représentés les spectres ^{13}C -HSQC des fractions d'HS issus des cellules MEF sauvages et des cellules MEF *ndst-1^{-/-}*. Les spectres sont globalement équivalents sur le plan de la qualité spectrale. En particulier, ils ne présentent pas de contamination notable par des acides nucléiques et permettent tous les deux une analyse quantitative d'un certain nombre de caractéristiques structurales des HS. Un premier examen montre pour les cellules MEF *ndst-1^{-/-}* la diminution de certains des pics associés aux espèces sulfatées.

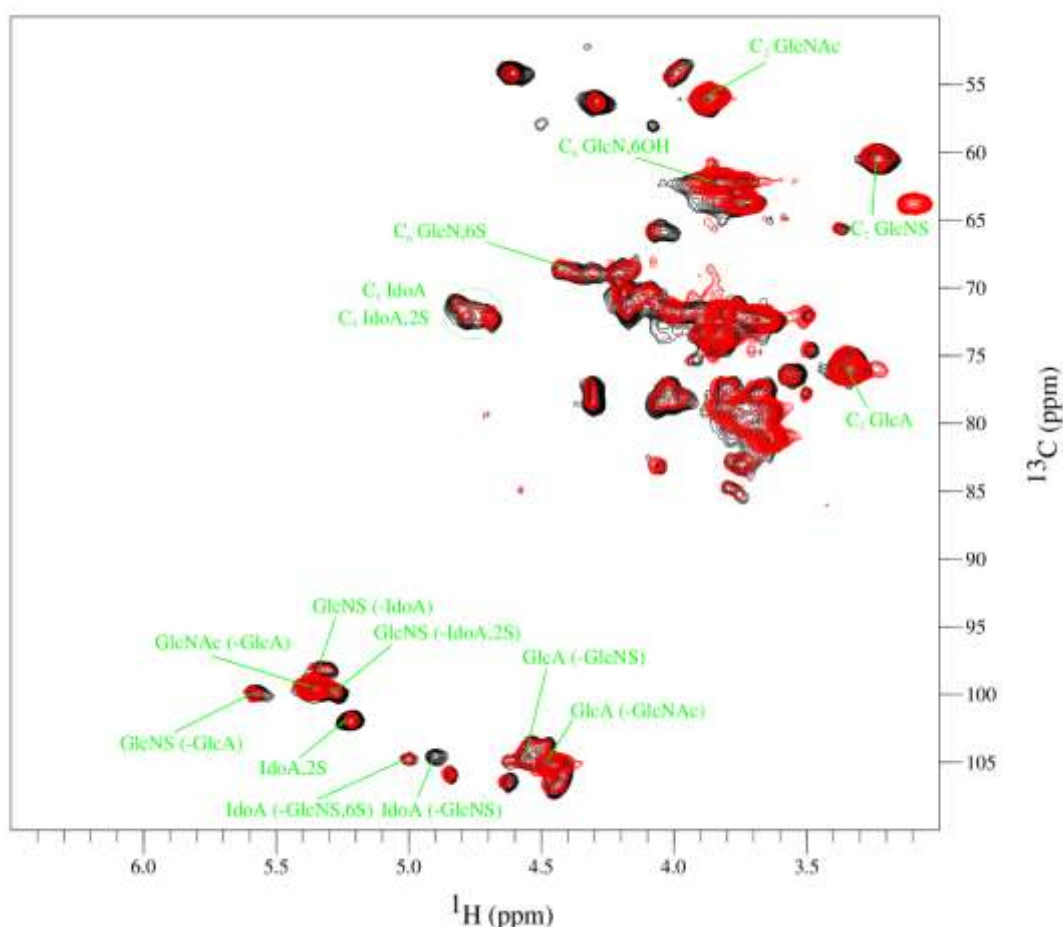


Figure 99 : Spectres RMN ^{13}C -HSQC d'HS cellulaires purifiés à partir de cellules MEF WT (noir) et *ndst-1*^{-/-} (rouge). Les échantillons sont analysés à 37°C pendant 16h dans un spectromètre 600MHz (spectre rouge) ou 700MHz (spectre noir).

L'intégration des pics d'intérêt pour l'analyse quantitative est faite au niveau des mêmes pics utilisés pour la détermination des caractéristiques structurales des HS issus des cellules HeLa.

La quantification par RMN révèle effectivement une baisse de la *N*-sulfatation générale des HS de MEF KO (34%) comparés à ceux des MEF WT (51%). Une baisse de l'épimérisation est également observée, passant de 36% pour les MEF WT à 14% pour les MEF KO. Cette réduction du taux d'épimérisation est effectivement en accord avec le mécanisme catalytique admis pour la C₅-épimérase où le substrat est un enchaînement GlcNS-GlcA ; la réduction de la quantité de substrat entrainant une baisse de la quantité en IdoA. En conséquence, la quantité de 2-*O*-sulfatation (IdoA,2S) est également diminuée dans les HS des MEF KO (9%) par rapport aux MEF WT (24%). Toutefois, l'efficacité de la HS2ST envers les IdoA n'est pas altérée car leur taux de 2-*O*-sulfatation est le même (~66% des IdoA sont 2-*O*-sulfatés). La 6-*O*-sulfatation des glucosamines ne varie pas dans les chaînes d'HS selon que la cellule exprime la NDST-1 (20%) ou pas (18%) (Figure 99). Le nombre de sulfate moyen par disaccharide diminue de 0,95 (pour les cellules WT) à 0,61 pour les cellules *ndst-1*^{-/-}.

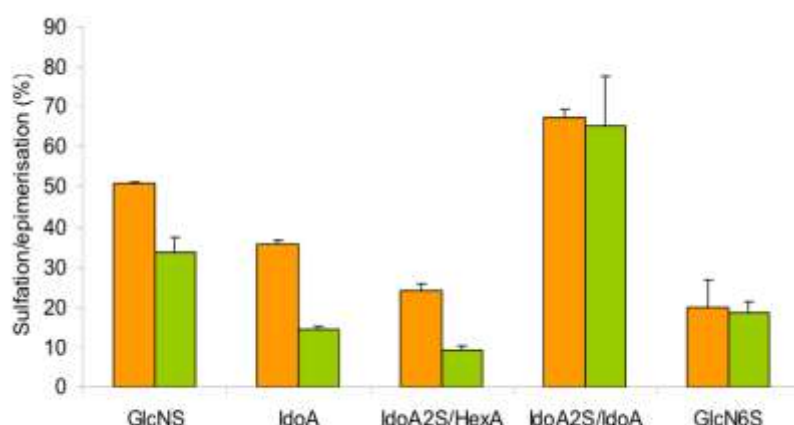


Figure 100 : Comparaison des profils de sulfatation et d'épimérisation des HS issus de cellules MEF WT (orange) et *ndst-1*^{-/-} (vert). Quantification d'après les résultats RMN.

L'analyse par RMN de la composition en disaccharide révèle que les MEF KO synthétisent des chaînes d'HS possédant une quantité en GlcNS-GlcA plus importante (51%) que les MEF WT (29%). Les glucosamines liées aux IdoA sont également différemment modifiées selon les deux souches cellulaires. Les WT forment principalement des IdoA reliés à des GlcNS (65%) alors que sur les cellules KO à la *ndst-1* les IdoA sont surtout reliés à des GlcNS,6S (88%). En revanche, pour les disaccharides contenant un GlcA, quel que soit l'état de la cellule, les GlcA-GlcNAc sont les disaccharides les plus nombreux comparés aux GlcA-GlcNS(±6S) (Figure 101).

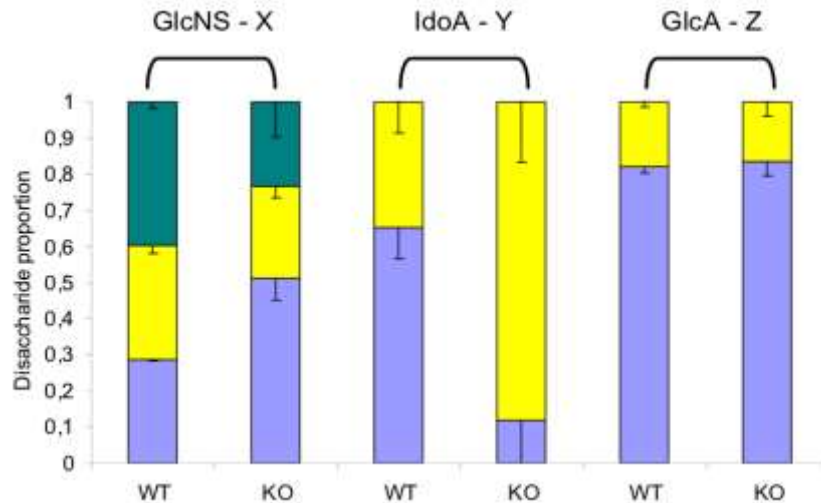


Figure 101 : Comparaison des disaccharides d'HS issus de cellules MEF WT (WT) et *ndst-1*^{-/-} (KO). Proportion en disaccharides d'HS de cellules MEF WT et *ndst-1*^{-/-} dans lesquels une GlcNS(±6S) est associée à X = GlcA (bleu), IdoA (jaune) ou IdoA,2S (vert) ; un IdoA à Y = GlcNS (bleu) ou GlcNS,6S (jaune) ; un GlcA est lié à Z = GlcNAc (bleu) ou GlcNS (jaune).

L'estimation de la longueur des chaînes entre les différentes cellules d'après le rapport de corrélations H₂-C₂ des glucosamines et H₅-C₅ du xylose (Figure 102) montre une longueur similaire d'une trentaine d'unités disaccharidiques pour les deux types cellulaires, ce qui correspond à une masse moléculaire d'une vingtaine de kilodaltons.

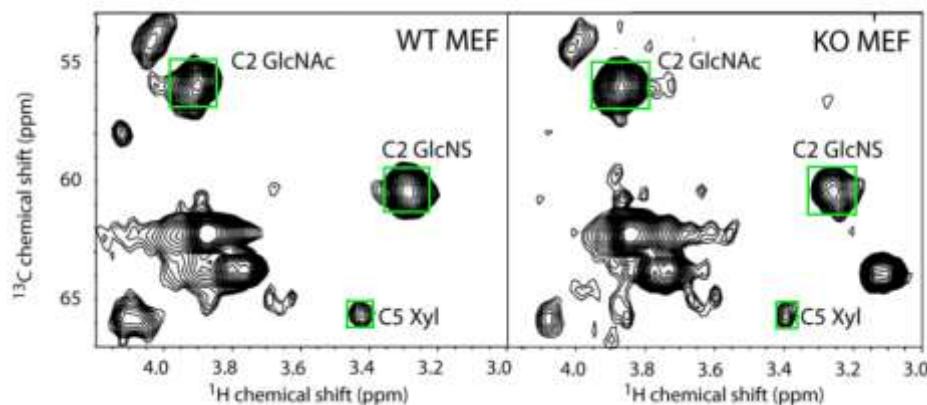


Figure 102 : Spectre RMN ¹³C-HSQC des déplacements chimiques H₂-C₂ des glucosamines et d'une paire H₅-C₅ du xylose d'HS de cellules MEF WT (gauche) et *ndst-1*^{-/-} (droite). Les encadrés verts représentent les bornes utilisées pour l'intégration des pics.

Parallèlement à cette étude, une détermination de la structure des HS par digestion puis analyse chromatographique de la composition en disaccharides a été menée par nos collaborateurs du groupe de Lena Kjellén qui nous a fourni les cellules MEF WT et KO. Les résultats obtenus sont en excellent accord avec les nôtres (voir publication Pegéot et al 2014 présentée en annexe 7).

L'utilisation de la RMN comme outil pour mesurer et analyser l'impact de l'absence d'une enzyme de biosynthèse des HS sur la structure même du polysaccharide est validée par la capacité à pouvoir déterminer les différences de taux de modifications des chaînes issues des MEF WT et des MEF *ndst-1*^{-/-} ainsi que leur structure disaccharidique.

Le protocole ainsi établi sur la base des cellules HeLa, et validé pour les cellules MEF, a été transposé aux cellules épithéliales Caco-2 pour l'étude de l'évolution du profil disaccharidique de ces cellules au cours d'un phénomène de différenciation.

2. Les cellules Caco-2

Les cellules Caco-2 sont issues de carcinome de cellules de côlon. Ces cellules ont été découvertes en 1977 (Fogh et al. 1977) et ont la propriété de se différencier spontanément après plusieurs jours de confluence dans les mêmes conditions de culture. Les cellules adoptent alors un phénotype plus proche des entérocytes (cellules de l'intestin grêle) avec la formation de microvillosités, de jonctions serrées, de transports vectoriels, ainsi qu'une expression différentielle en enzymes (Rousset 1986)... De plus, dans ces cellules, en fonction du stade de différenciation, la composition en disaccharides des HS varie (Salmivirta et al. 1998). Les cellules Caco-2 apparaissent donc comme un modèle cellulaire intéressant pour appliquer notre protocole d'analyse et l'étendre à l'étude de l'évolution de la structure des HS au cours de la différenciation cellulaire.

a) Détermination par cytométrie en flux de la présence d'HS au niveau des surfaces cellulaires

La première étape, avant d'appliquer le protocole mis en place et détaillé dans les parties précédentes, a été de déterminer la présence d'HS au niveau des surfaces cellulaires. Pour ce faire, des analyses de cytométrie en flux ont été réalisées après préparation des cellules Caco-2 non-différenciées et différenciées incubées avec différents anticorps dirigés spécifiquement contre les HS ou les CS.

Par rapport au contrôle isotypique, les cellules non-différenciées incubées avec les anticorps primaires dirigés contre les HS et secondaires présentent une intensité de fluorescence importante. Les cellules marquées aux anticorps dirigés contre les HS ont une répartition de l'intensité de fluorescence médiane autour de 40 (courbe rouge), soit une intensité dix fois plus importante par rapport au contrôle isotypique (médiane de la fluorescence à 4, courbe bleue) (Figure 103A). Les cellules incubées avec les anticorps dirigés contre les CS, quant à elles, montrent une intensité de fluorescence qui est similaire au contrôle isotypique (courbe verte). Ces résultats indiquent d'une part la présence d'HS à la surface des cellules en quantité notable, et d'autre part que les CS sont, pratiquement indétectables.

Comparé aux cellules HeLa, l'intensité de fluorescence due à l'anticorps 10E4 pour les cellules Caco-2 non-différenciées est quatre fois moins importante (courbe rouge Figure 68). Cela peut s'expliquer soit par un nombre plus faible de chaînes présentes à la surface des cellules CaCo-2 et/ou par une structure de chaînes différentes, possédant en plus faible quantité le nombre d'épitope de reconnaissance de l'anticorps.

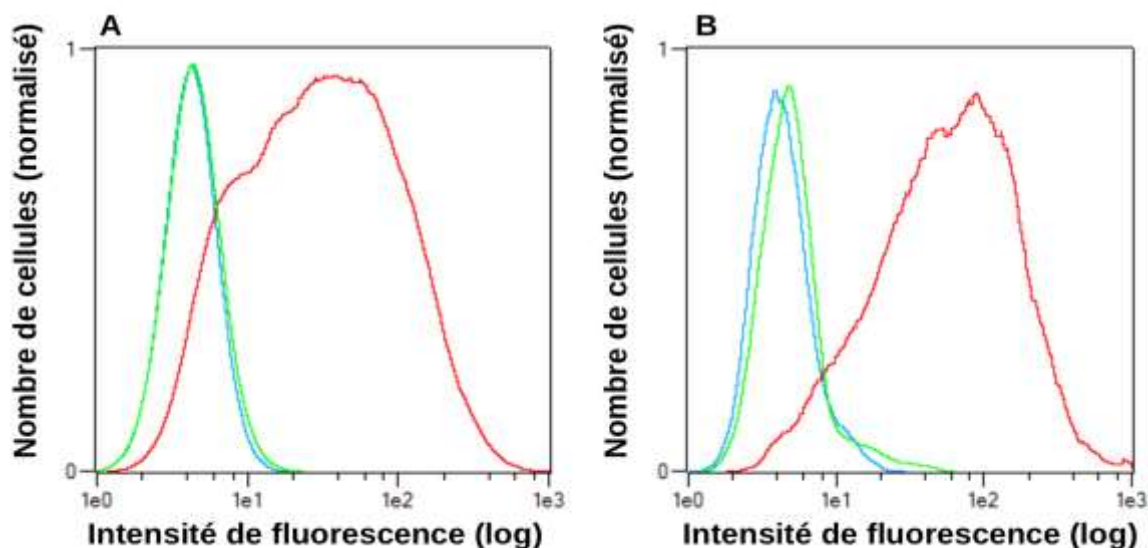


Figure 103 : Analyse par cytométrie en flux des cellules Caco-2 avant différenciation (A) et après différenciation (B). Des anticorps spécifiques sont utilisés pour détecter les GAG à la surface des cellules. Détection des HS (rouge) et CS (vert), contrôle isotypique (bleu).

Après différenciation, les cellules Caco-2 incubées avec les anticorps de reconnaissance des HS montrent une intensité de fluorescence plus importante que pour les cellules non-différenciées (médiane autour de 80). Cela représente une fluorescence deux fois plus intense par rapport aux Caco-2 non-différenciées (Figure 103B). Toutefois, le marquage semble peu uniforme ce qui se traduit par une largeur de pic importante au niveau de l'intensité de fluorescence. Cela peut s'expliquer par l'hétérogénéité des cellules et de leur expression en HS comme par des difficultés expérimentales liées à la nature même de ces cellules. En effet, celles-ci ayant tendance à former des jonctions serrées, même après traitement à l'acutase, elles peuvent toujours être en forte interaction et ne pas passer individuellement au niveau du laser, introduisant un biais au niveau de la détection de la quantité de fluorescence. Cette hétérogénéité est également observée pour les cellules Caco-2 non-différenciées.

Le marquage des cellules différenciées avec les anticorps dirigés contre les CS montrent un léger déplacement du pic de fluorescence par rapport au contrôle isotypique (Figure 103B). Cela peut suggérer une expression accrue des PG cellulaires porteurs de chaînes CS, comme une modification des structures des CS après différenciation présentant en plus grand nombre l'épitope de reconnaissance spécifique de l'anticorps CS-56.

Les analyses de cytométrie en flux des cellules Caco-2 montrent donc avant et après différenciation la présence effective d'HS à leur surface. Pour autant, à partir de ces données, il n'est pas possible de conclure à des différences significatives dans l'expression en GAG au cours de la différenciation ; l'augmentation de l'intensité de fluorescence pouvant tout aussi bien s'expliquer par des biais expérimentaux liés à la nature des cellules.

En conclusion, compte tenu de la présence des HS sur les Caco-2 en quantité relativement suffisante, le protocole mis en place pour les cellules HeLa semble donc directement transposable aux cellules Caco-2, et ce, pour le même nombre de boîtes cultivées. Cela devrait également permettre une approche pour déterminer l'évolution du profil saccharidique au cours de la différenciation.

par les pics dans la région 6ppm ^1H , 90 ppm ^{13}C) mais qui n'altère pas la détermination des modifications des HS. La quantification des caractéristiques structurales des HS des cellules Caco-2 a été faite de la même façon que pour les HS des cellules HeLa en utilisant les mêmes corrélations.

Le profil saccharidique moyen des HS obtenu à partir des cellules Caco-2 non-différenciées (Figure 105, graphes orange) est similaire à celui des HS de cellules HeLa (Figure 93). Sur 100 disaccharides, 48 en moyenne possèdent une GlcNS pour 46 dans les HeLa. De ces glucosamines, 22 sont 6-*O*-sulfatées contre 23 pour les HeLa. Le taux d'épimérisation des acides uroniques en acides iduroniques est de 27%, semblable à celui des cellules HeLa (29%). Le taux de 2-*O*-sulfatation est également équivalent : 16% pour les Caco-2 non-différenciées et 19% pour les cellules HeLa. La sulfatation est en moyenne par disaccharide de 0,83 ce qui est peu différent de celle des cellules HeLa qui est de 0,88.

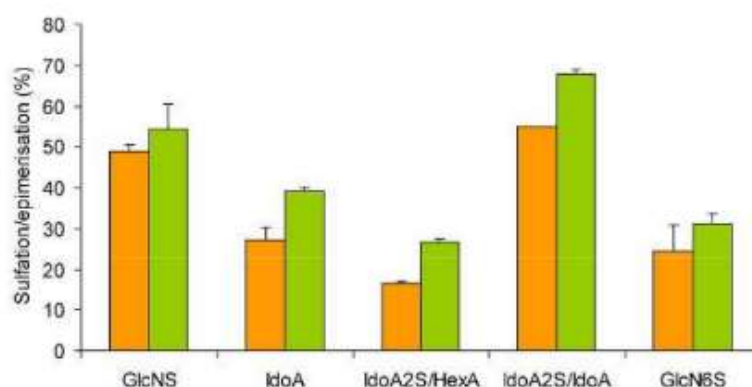


Figure 105 : Comparaison des profils de sulfatation et d'épimérisation des HS issus de cellules Caco-2 non-différenciées (orange) et différenciées (vert). Quantification d'après les résultats RMN.

L'estimation de la longueur de chaîne est par contre difficile pour cet échantillon. Le signal du xylose est en effet très peu intense comparé aux autres signaux. Cela suggère que les chaînes sont longues, au moins de la taille de celles des HS des cellules HeLa, mais il est difficile de donner une valeur du nombre de disaccharide dans ce cas (Figure 106).

Après différenciation, le taux de modification augmente. 54% de glucosamines sont *N*-sulfatées et 31% 6-*O*-sulfatées (contre 48% et 22% respectivement pour les HS de Caco-2 non-différenciées). L'épimérisation des acides glucuroniques en iduroniques passe de 29 à 39% après différenciation dont plus des deux tiers sont 2-*O*-sulfatés (un peu plus d'un sur deux dans les HS de Caco-2 non-différenciées) (Figure 105). Le nombre moyen de sulfate par disaccharide est de 1,1, soit une augmentation d'un tiers de la sulfatation moyenne par rapport aux HS des cellules non-différenciées. Le signal du xylose est également plus intense et le ratio GlcN/Xyl de l'ordre de la trentaine, soit des chaînes au moins deux à trois fois plus courtes (Figure 106).

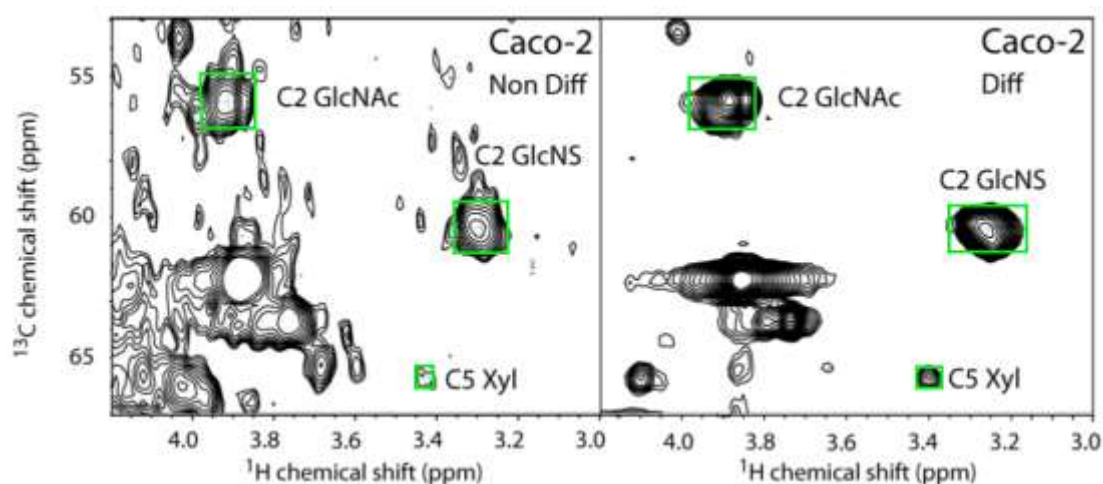


Figure 106 : Spectre RMN ^{13}C -HSQC des déplacements chimiques $\text{H}_2\text{-C}_2$ des glucosamines et d'une paire $\text{H}_5\text{-C}_5$ du xylose d'HS de cellules Caco-2 non-différenciées (gauche) et différenciées (droite). Les encadrés verts représentent les bornes utilisées pour l'intégration des pics.

L'analyse de l'ose voisin montre une augmentation de la quantité de GlcNS($\pm 6\text{S}$)-IdoA,2S après différenciation avec 40% de probabilité d'avoir un IdoA,2S après une glucosamine *N*-sulfatée contre 21% dans les HS de Caco-2 non-différenciées. Cela est corrélé avec la baisse de probabilité d'avoir un GlcA après une glucosamine *N*-sulfatée, proportion en GlcNS($\pm 6\text{S}$)-GlcA qui passe de 52% dans les HS de Caco-2 non-différenciées à 33% dans les cellules différenciées (Figure 107).

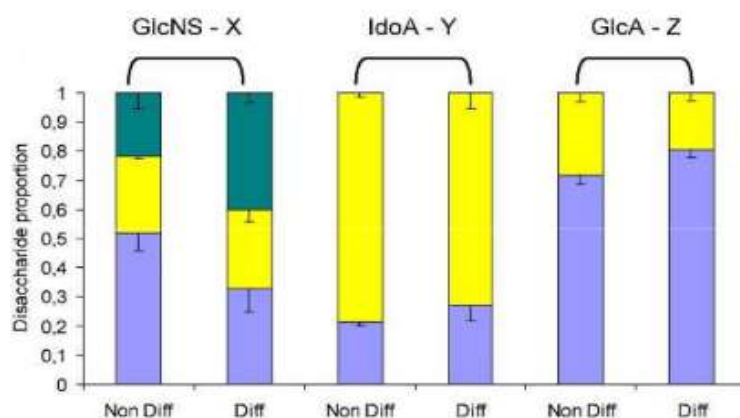


Figure 107 : Comparaison des disaccharides d'HS issus de cellules Caco-2 non-différenciées (Non Diff) et différenciées (Diff). Proportion en disaccharides d'HS de cellules Caco-2 non-différenciées et différenciées dans lesquels une GlcNS($\pm 6\text{S}$) est associée à X = GlcA (bleu), IdoA (jaune) ou IdoA,2S (vert) ; une IdoA à Y = GlcNS (bleu) ou GlcNS,6S (jaune) ; un GlcA est lié à Z = GlcNAc (bleu) ou GlcNS (jaune).

En revanche, les IdoA sont majoritairement suivis de GlcNS,6S et les GlcA de GlcNAc quel que soit l'état de différenciation, et dans des valeurs similaires ($\approx 80\%$ pour chaque).

Ainsi, notre protocole d'analyse nous a permis de suivre les modifications induites par la différenciation cellulaire dans la structure des HS cellulaires issues de cellules Caco-2. Une augmentation de la *N*-sulfatation, 2-*O*-sulfatation et 6-*O*-sulfatation et donc de la sulfatation globale est observée dans notre étude par RMN ce qui est en désaccord avec l'analyse

disaccharidique réalisée par Salmivirta *et al.* 1998 dont les chaînes d'HS sont, certes plus 2-O-sulfatées mais avec une N-sulfatation constante et même une baisse de la quantité de 6-O-sulfates après différenciation des cellules Caco-2, résultant en une sulfatation globale stable. Les taux de N-sulfatation des HS avant et après différenciation des cellules obtenus (48 % et 54 % respectivement) sont en accord avec les analyses saccharidiques de Lévy *et al.* 1988 (50% et 55% pour les HS des cellules non-différenciées et différenciées respectivement) (Levy *et al.* 1988).

Concernant l'épimérisation, notre étude montre un taux d'épimérisation qui passe de 30 % à 39 % dans les chaînes d'HS issues de cellules non-différenciées et différenciées respectivement, ce qui est en accord avec Salmivirta *et al.* 1998 (30 et 36 % respectivement).

Cependant, pour les longueurs moyennes des chaînes, nos observations sont en désaccord avec les résultats de la littérature. Nous observons par analyse RMN, un raccourcissement approximativement de moitié des chaînes. L'étude de Salmavirta concluait plutôt à une stabilité de celles-ci au cours du processus de différenciation alors que celle de Lévy concluait à une élongation. Les divergences entre les résultats de la littérature ainsi qu'avec ceux obtenus par analyse RMN montrent la complexité de l'étude des HS. Ces différences observées peuvent à la fois provenir des différentes conditions de culture cellulaire entre les études, du marquage et de la préparation par dépolymérisation des HS ainsi que des techniques d'analyse.

En conclusion, ces résultats montrent que les caractéristiques des HS issus de différentes lignées cellulaires peuvent être déterminées grâce à une analyse par RMN. Cette analyse non-destructive pour les polysaccharides est menée sur un matériel directement issu des cellules. Le choix de la RMN comme méthode fiable pour l'analyse du profil saccharidique a été précédemment validé par l'étude des HS issus des MEF en accord avec l'analyse disaccharidique et a fait l'objet d'une publication (annexe 7) (Pegeot *et al.* 2014).

D. Etude des HS dans le contexte du protéoglycane

L'étude des HS par RMN liquide s'est donc révélée possible sur les HS entiers marqués issus de surfaces cellulaires. Elle a permis de déterminer le profil saccharidique et de modifications de ceux-ci à partir de quantités de cellules abordables et de temps d'analyse raisonnables. Toutefois le fait d'extraire les HS conduit à perdre des informations de localisation et de coopération éventuelle des chaînes dans le contexte cellulaire. L'obtention de ces quantités d'HS par RMN du liquide nous a fait envisager la possibilité d'une étude directe sur des cellules intègres par des techniques de RMN solide (ou ssNMR pour *solid state NMR*). Cela a été mené sur le modèle des cellules épithéliales HeLa.

1. Etudes des sucres à partir de cellules entières

Les cellules HeLa analysées ont été cultivées comme précédemment dans un milieu marqué (^{13}C -glucose) et ont été récupérées après détachement du support par une solution d'EDTA (versène) afin de préserver les molécules présentes sur les surfaces cellulaires. Les

cellules sont, après lavage, directement introduites dans un rotor pour leur analyse par RMN solide par centrifugation.

Dans la Figure 108 sont présentés les résultats d'expériences préliminaires faites avant le début de ce travail. Le spectre ^{13}C 1D en démarrant par une excitation directe du ^{13}C et obtenu sur des cellules HeLa entières est présenté en bleu dans cette Figure 108. Dans ce spectre montrant toutes les résonances carbones détectables, on peut immédiatement observer la présence de signaux résolus nettement au-dessus du bruit de fond. De plus, une partie de ces signaux correspond aux fréquences où sont attendus les signaux des espèces glycosidiques (entre 50ppm et 100ppm). On observe également des signaux aux fréquences des carbones des fonctions carboxyliques (172 ppm) comme aux carbones des chaînes insaturées (128 ppm). De même, un certain nombre de signaux résolus sont présents vers 20 ppm, région où sont attendus les signaux des groupements *N*-acétyles des glucosamines non-modifiées. Ces premiers résultats ont suggéré la présence d'un ensemble d'espèces de nature glycosidique.

Des cellules HeLa ont été ensuite traitées aux héparinases I et III. Dans ce cas, on peut noter qu'après digestion des surfaces cellulaires, les intensités de certains pics diminuent, en particulier à 100ppm dans la région des carbones anomériques des sucres, ce qui suggère qu'une partie de ces signaux pourrait être attribuée à des espèces de type HS (spectre rouge).

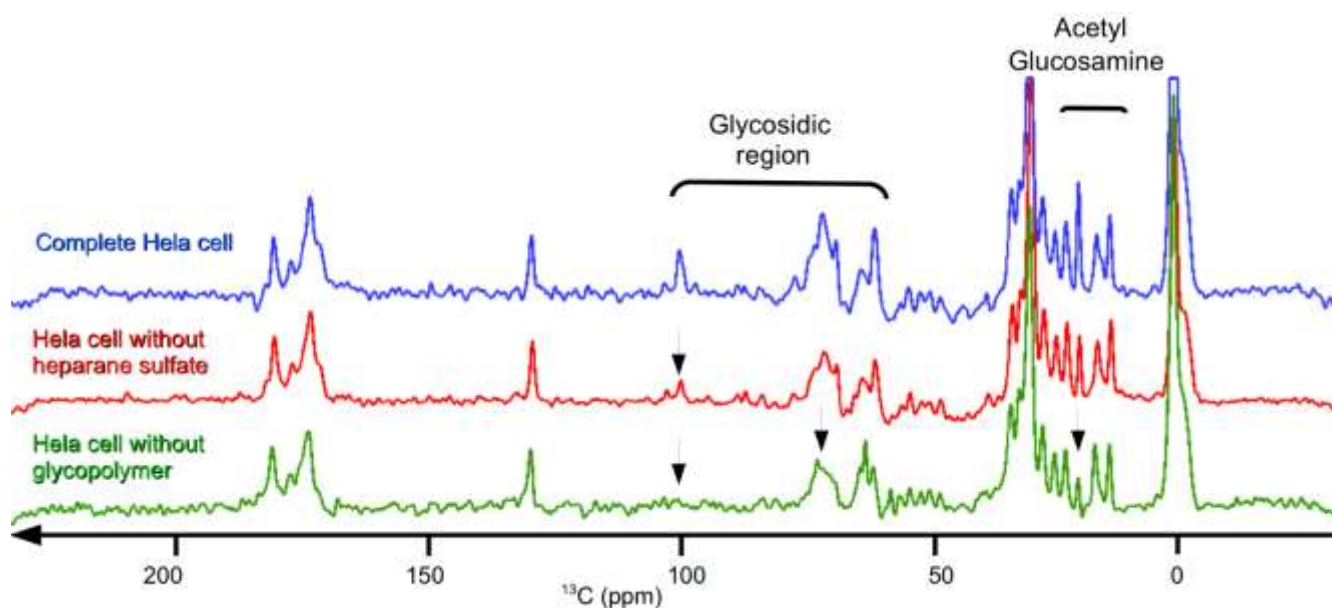


Figure 108 : Analyses par RMN solide d'un spectre ^{13}C après excitation directe du carbone. Détection enregistrée sur des cellules HeLa intactes (bleu), traitées aux héparinases I et III (rouge) et traitées aux héparinases I et III, chondroïtinase ABC et *N*-glycosidase (vert). Les spectres sont enregistrés à 12°C pendant 30 minutes à une vitesse de rotation à l'angle magique (MAS) de 12kHz. Les flèches noires indiquent les signaux affectés par les traitements enzymatiques.

Enfin, lorsque les cellules sont traitées simultanément aux héparinases, chondroïtinase et *N*-glycosidase, les signaux à 100ppm sont pratiquement abolis, les signaux vers 80-60ppm sont fortement diminués comme également un pic à 23ppm alors que les autres signaux ne sont pas touchés. Les cellules étant intactes, cela suggère fortement que les espèces à l'origine

de ces signaux sont accessibles aux traitements enzymatiques et donc exposées à la surface de la cellule.

Tous ces résultats tendent à montrer qu'il existe à la surface cellulaire une quantité d'espèces glycosidiques suffisamment importante pour être observable par RMN du solide. De plus, le fait qu'une partie de ces espèces soit sensible à l'action des héparinases suppose une contribution significative des HS au niveau des membranes.

Ces résultats étant alors très encourageants et comme la technique de RMN 1D ne permet pas une caractérisation plus précise de la nature de ces composés, des expériences bidimensionnelles ont été réalisées afin d'améliorer la résolution et de caractériser plus finement les espèces présentes.

De même afin d'obtenir la plus grande sensibilité des analyses, la totalité des 200 millions de cellules HeLa récoltées ont été placées dans un rotor à large volume (7mm de diamètre). Les analyses ont été effectuées à l'Institut Nanoscience et Cryogénie (INAC) sur un spectromètre 400MHz équipé d'une sonde solide (collaboration avec le Dr. Hediger). L'échantillon étant plus conséquent, une meilleure sensibilité est attendue pour ces expériences bidimensionnelles.

Des expériences de corrélation ^{13}C - ^{13}C ont été réalisées permettant d'observer des corrélations aux fréquences de carbones vicinaux (Figure 109). Sur la figure est représenté le spectre obtenu. Malgré la quantité de cellules utilisées, la sensibilité du spectre n'est pas optimale. On peut toutefois observer dans la région des carbones anomériques (90-105ppm, encadré vert), plusieurs signaux diagonaux correspondant à autant de carbones anomériques. Chacun de ces carbones est voisin d'un autre carbone dont le déplacement chimique est dans la gamme des 70ppm. De même, des corrélations diagonales et hors diagonales sont observées entre carbone dont les déplacements chimiques sont compris entre 60 et 80 ppm. Cela est en accord avec ce qui est attendu pour les déplacements chimiques des oses. Par contre, nous n'observons pas de corrélation C_1 - C_2 aux déplacements chimiques attendus pour des GlcA (105ppm-76ppm), des GlcNAc (100ppm-56ppm) et GlcNS (100ppm-60ppm).

Il est ainsi possible d'observer les signaux de sucres pour des cellules eucaryotes HeLa entières par ssNMR. Cette technique est suffisamment résolutive pour déterminer certaines corrélations. Cependant, bien que des HS puissent être présents au niveau des surfaces cellulaires (cf traitement aux héparinases Figure 108), nous n'avons pas pu observer leurs déplacements chimiques par les expériences 2D de corrélation ^{13}C - ^{13}C par le couplage scalaire (couplage *via* les liaisons chimiques). Cela pourrait s'expliquer par une trop faible quantité d'HS pour les analyses dans ces conditions et donc une trop faible sensibilité. Des expériences dites par *Cross Polarization* (CP) et démarrant de l'aimantation du proton ont été testées. Malgré le gain de sensibilité escompté du fait du rapport gyromagnétique du proton, ces expériences n'ont pas permis d'augmenter l'aimantation détectée. Cette faible efficacité du transfert CP entre le proton et le carbone peut s'expliquer par la présence d'une flexibilité importante au sein de la molécule et qui tend à réduire le couplage dipolaire ^1H - ^{13}C (résultat non-montré).

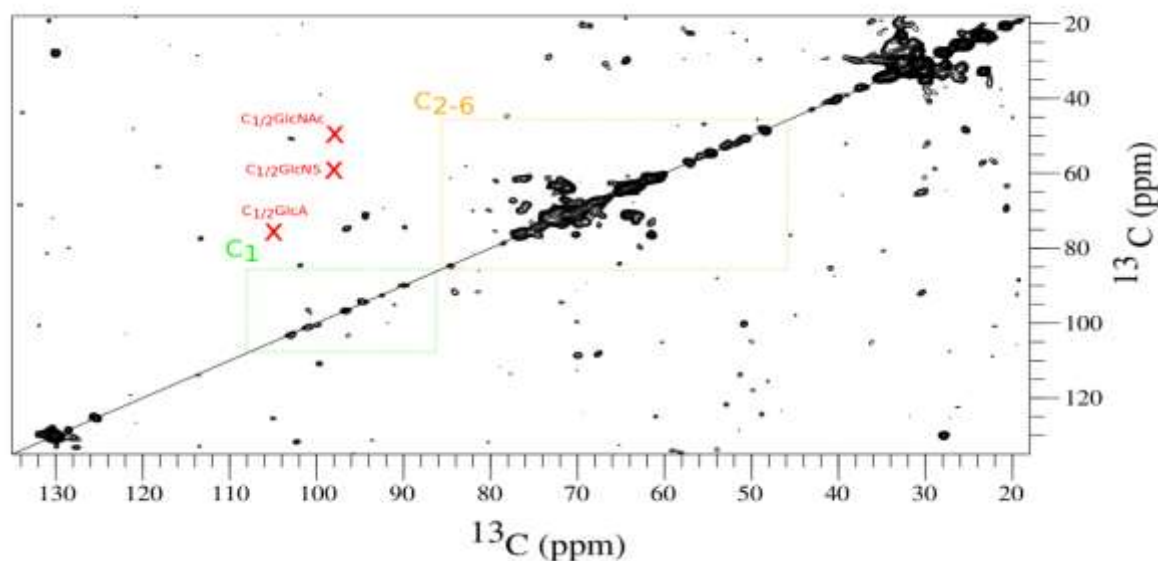


Figure 109 : Expérience par ssNMR de corrélation ^{13}C - ^{13}C après excitation directe du ^{13}C et en utilisant le couplage scalaire lors du transfert ^{13}C - ^{13}C . Ce spectre a été enregistré sur cellules entières HeLa à 4°C, avec un temps expérimental de 15h et à une vitesse de rotation à l'angle magique de 3,4 kHz. L'encadré vert correspond aux corrélations attendues pour les carbones anomériques et l'encadré orange aux autres corrélations attendues pour les GAG. Les corrélations C_1 - C_2 attendues pour les GlcNAc, GlcNS et GlcA sont indiquées par une croix rouge.

Dans le but d'augmenter significativement la sensibilité de l'expérience, des techniques de polarisation nucléaire dynamique (DNP) ont été testées sur les cellules HeLa.

2. Premières analyses par polarisation nucléaire dynamique (DNP)

Afin d'augmenter la sensibilité dans le cas de matériaux poreux ou de l'étude de surfaces, des expériences de DNP ont été proposées. Il s'agit de créer une aimantation au niveau d'un électron célibataire porté par un radical et de transférer l'aimantation vers les noyaux du matériel étudié. Pour ce faire, un radical, ou agent polarisant, tel que le totapol (1-(TEMPO-4-oxy)-3-(TEMPO-4-amino)propan-2-ol) est mis en contact avec le matériau et est ensuite excité par une émission micro-onde (Song et al. 2006). Cette expérience a été testée sur des cellules bactériennes (Takahashi et al. 2013) et a montré qu'il était possible d'augmenter sensiblement les signaux des molécules se trouvant à la surface cellulaire. Nous avons donc cherché à utiliser un protocole très similaire afin d'observer les chaînes de sucres se trouvant à la surface des cellules épithéliales. Après avoir mis en contact le totapol et une matrice contenant du DMSO, nous avons excité à l'aide d'un gyrotron l'électron du totapol et enregistré des expériences de type DQ (double quanta). Dans ces expériences, on trouve en abscisse les déplacements chimiques directement observés et en ordonnée la somme des deux déplacements chimiques des carbones voisins. Ainsi, une corrélation est observée entre deux carbones vers 100ppm et 58ppm (en abscisse) donnant un signal vers 158 (100+58ppm) en ordonnée (Figure 110). Ces déplacements chimiques sont similaires à ceux attendus pour les C_1 et C_2 des GlcNAc. Un signal faible est également observé vers 105 et 80ppm (abscisse) et 185ppm (ordonnée). Ces déplacements chimiques semblent correspondre à ceux attendus pour les C_1 et C_2 des GlcA. Toutefois la sensibilité, malgré la technique, reste très limitée et ne

permet pas d'affirmer avec certitude la présence de GlcA bien que cette observation soit encourageante. En particulier, il a été montré que la sensibilité est très dépendante de la nature du matériel étudié et du mode de préparation de l'échantillon (quantité de radical, nature de la matrice, méthode de trempage...). Ces premiers essais restent donc une voie à explorer et à optimiser afin de caractériser finement les HS par ssNMR.

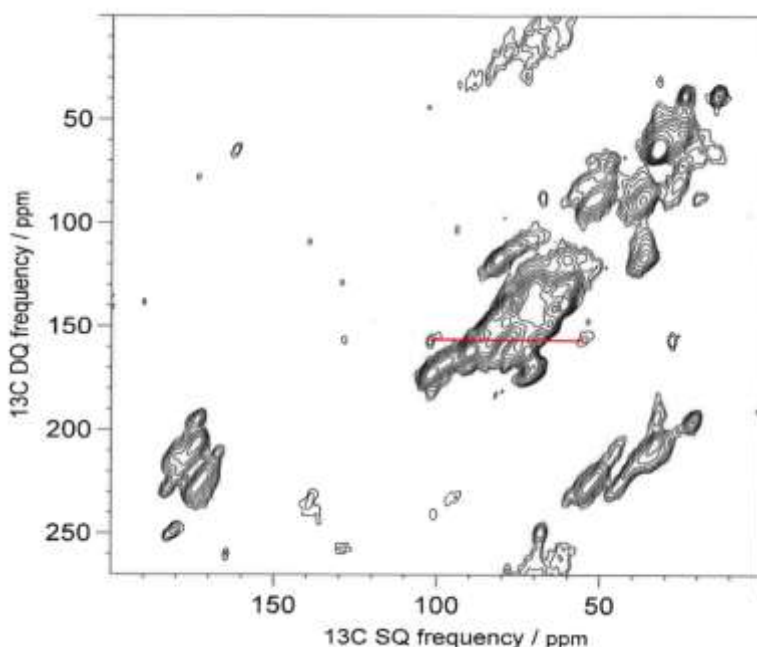


Figure 110 : Expérience Double Quanta (DQ) ssNMR par DNP de cellules HeLa entières marquées. Le spectre est enregistré sur des cellules HeLa à 4°C pendant 15h en rotation à l'angle magique.

Le manque de sensibilité des expériences peut s'expliquer, en plus de la technique RMN en elle-même, par la trop faible quantité en matériel d'intérêt au sein des rotors d'analyse. En effet, le volume occupé dans le rotor est principalement celui des organites de la cellule (cytosol, noyau...) et peu par les membranes comprenant les HSPG. Un moyen donc d'augmenter la concentration en HS est d'étudier alors les membranes plasmiques isolées.

3. Etudes des sucres dans les membranes plasmiques isolées de cellules HeLa

L'analyse des résultats obtenus sur les cellules entières est limitée par la sensibilité des expériences. Afin d'augmenter au maximum la quantité d'échantillon analysée, nous avons utilisé des préparations contenant uniquement des membranes cellulaires. En éliminant les différents organites qui occupent un volume conséquent, nous espérons occuper au maximum le volume du rotor avec les molécules d'intérêt.

Une fois préparés, nous avons analysé ces échantillons par différents types d'expériences RMN. En premier lieu, nous avons enregistré des expériences ^{13}C en excitation directe qui permettent d'observer les signaux de toutes les espèces contenant du ^{13}C dans l'échantillon indépendamment de leur flexibilité. Le spectre, comme l'illustre la figure suivante, ne montre qu'un nombre limité de résonances à des déplacements chimiques similaires à ceux du spectre 1D carbone des cellules entières. Nous observons plusieurs pics

(93, 97 et 100ppm) dans la région des carbones anomériques dont le plus intense est à 100ppm (Figure 111A) qui est le déplacement chimique attendu pour les C₁ des glucosamines. De même, on observe un massif de pics entre 60 et 80 ppm. Ces signaux sont dans la zone attendue pour les C₂-C₆ des hexoses. Par contre, nous n'observons ni signal à 56ppm, ce qui est le déplacement chimique des C₂ des GlcNAc, représentant 55% des glucosamines selon les études par RMN liquide, ni de pic vers 105ppm où est attendu le signal des carbones anomériques des acides uroniques.

Une idée de la flexibilité des molécules présentes dans la préparation des membranes peut être obtenue par la mise en place d'expérience de polarisation croisée (CP) et INEPT (*insensitive nuclei enhanced by polarization transfer*). La première technique (CP) permet d'observer uniquement les parties suffisamment rigides pour donner lieu à un couplage dipolaire alors que la seconde n'utilise que le couplage scalaire efficace pour transférer l'aimantation que dans les parties plus flexibles. L'expérience INEPT révèle la présence de sucres flexibles mais les corrélations attendues pour les C₂-glucosamines (60 et 56ppm) et les C₁-uronates (105ppm) ne sont pas clairement observées (Figure 111B). De même pour les analyses par CP, les déplacements chimiques observés n'appartiennent vraisemblablement pas non plus aux HS (Figure 111C).

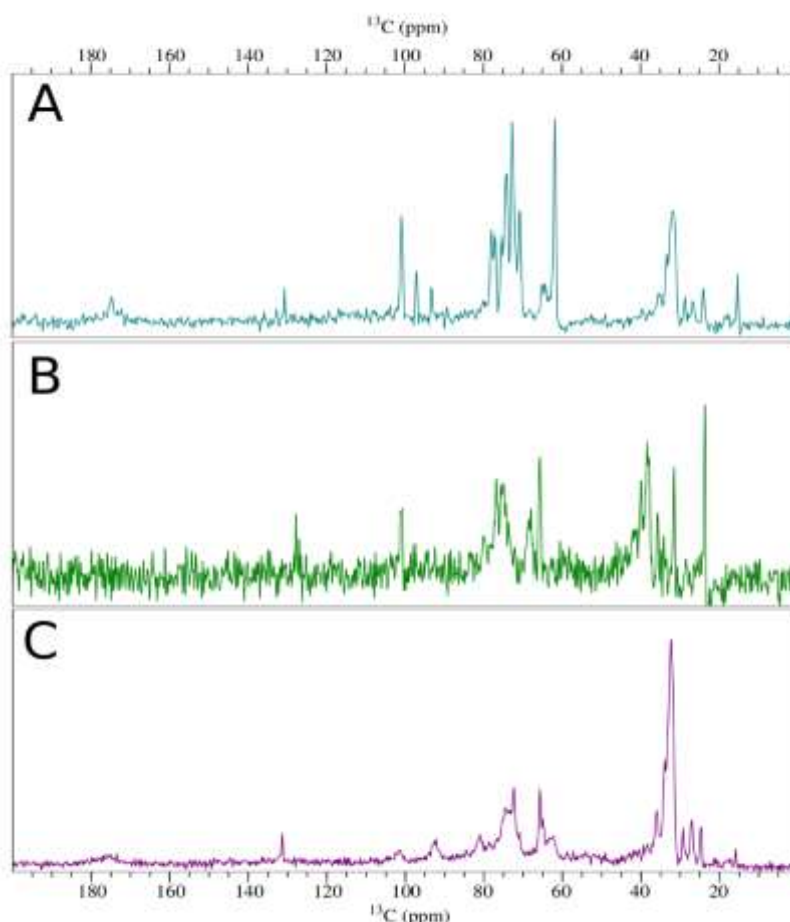


Figure 111 : Expériences par ssNMR de corrélations ¹³C par excitation directe (A), INEPT (B) et CP (C) d'une préparation de membranes de cellules HeLa. Les cellules sont analysées à 8°C à une vitesse de rotation de 12kHz pendant 11mn (¹³C-directe), 3mn (INEPT) ou 31mn (CP) dans un rotor 3,2mm.

En conclusion de ces expériences, comme dans les cellules entières, nous pouvons observer dans les membranes isolées des signaux correspondant très probablement à des sucres mais par contre il n'est pas possible de détecter la signature des HS.

Dans le but de déterminer de manière plus fine les espèces contenues dans l'échantillon, des expériences bidimensionnelles de corrélation ^{13}C - ^{13}C sont réalisées. Trois pics intenses à 93, 97 et 100ppm dans la zone des carbones anomériques sont observés et qui corrént avec leur carbone vicinal à 72, 75 et 72ppm respectivement (Figure 112). Dans la région entre 60ppm et 80ppm sont observées également des corrélations diagonales et hors-diagonales entre plusieurs carbones. Comme pour les cellules entières, les corrélations C_1 - C_2 attendues pour les HS (GlcA : 105-76ppm ; GlcNAc : 100-56ppm ; GlcNS 100-60ppm) ne correspondent pas à celles observées par RMN solide.

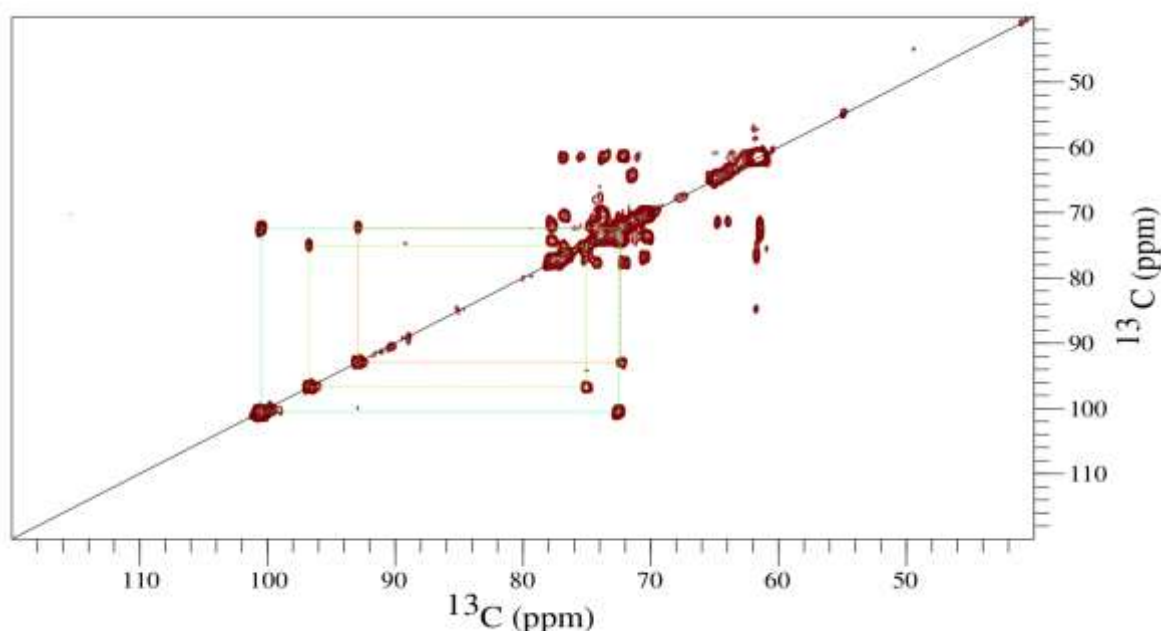


Figure 112 : Expérience ssNMR de corrélation ^{13}C - ^{13}C par couplage scalaire avec excitation directe du carbone enregistrée sur une préparation de membranes de cellules HeLa marquées. Température d'analyse : 8°C ; temps d'analyse : 37h ; vitesse de rotation : 12kHz.

Enfin, des expériences INEPT bidimensionnelles de corrélation ^1H - ^{13}C ont également été réalisées afin de déterminer les déplacements chimiques des protons liés directement aux ^{13}C observés dans les espèces relativement flexibles.

L'analyse du spectre obtenu (Figure 113) indique la présence des 3 carbones anomériques observés également par corrélation ^{13}C - ^{13}C . Le carbone observé à 100ppm est couplé à un carbone dont le déplacement chimique est de 5,5ppm, ce qui est similaire aux déplacements chimiques $^1\text{H}_1$ - $^{13}\text{C}_1$ des glucosamines. Toutefois, ce même carbone est couplé avec un carbone voisin dont le déplacement chimique est de 72ppm, et donc différent des déplacements chimiques C_2 des GlcNAc (56ppm) et GlcNS (60ppm) (Figure 112). Les autres déplacements chimiques des carbones anomériques observés sur les membranes (97 et 93ppm) ont des déplacements chimiques $^1\text{H}_1$ - $^{13}\text{C}_1$ différents des carbones anomériques

attendus pour les HS dont un spectre HSQC (en vert) est superposé au spectre 2D INEPT pour comparaison.

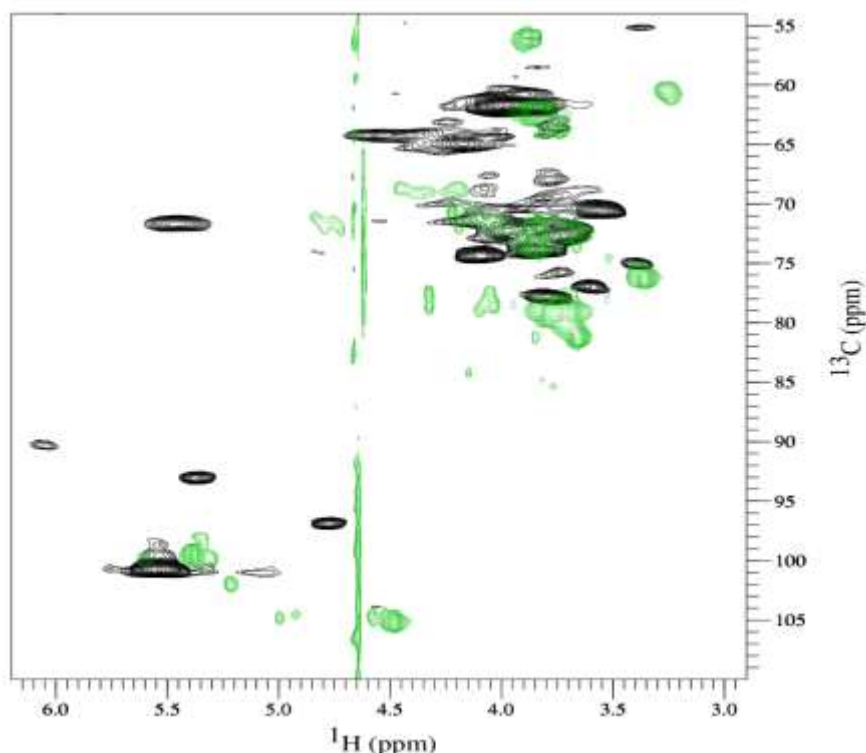


Figure 113 : Expérience ssNMR INEPT de corrélation ^1H - ^{13}C d'une préparation de membranes de cellules HeLa (noir) comparé à un spectre HSQC liquide d'HS commercial (vert). Temps d'analyse INEPT : 28h ; vitesse de rotation : 12kHz ; température d'analyse : 8°C.

A partir des préparations de membranes isolées d'HeLa, les analyses de spectroscopie RMN du solide de ces composés montrent la présence de différents sucres avec différentes flexibilités à leur surface. Des expériences bidimensionnelles permettent de déterminer les différents déplacements chimiques de ces espèces. Toutefois, ces déplacements chimiques ne correspondent pas à ceux attendus pour les HS et la nature de ces sucres reste encore à déterminer.

4. Etudes par RMN liquide des membranes de cellules HeLa solubilisées dans des détergents

Nous n'avons donc pas pu observer de signaux attribuables aux HS dans les préparations membranaires des cellules HeLa par RMN du solide bien que ces espèces soient présentes à la surface des cellules. Afin de comprendre si cela provient de la préparation de l'échantillon ou des conditions d'analyses, nous avons solubilisé les membranes préparées à partir des cellules HeLa et nous les avons analysées par RMN liquide.

Nous avons solubilisé les membranes grâce à trois détergents différents : le Triton X-100, l'octylPOE (octylpolyoxyéthylène) et le DPC (n-dodécylphosphocholine). Les résultats obtenus par RMN grâce à ces trois détergents présentent les mêmes caractéristiques. Seul le spectre enregistré sur des membranes solubilisées dans le détergent octylPOE est présenté ici

(Figure 114). Les spectres issus des membranes solubilisées au triton X-100 et au DPC sont présentés en annexe 6.

Nous retrouvons dans ce spectre pratiquement toutes les corrélations observées dans un spectre de type INEPT obtenu sur les fragments de membranes (Figure 114). Cela montre que les espèces que nous observons dans les membranes s'observent également en solution. De plus, il n'y a pas d'influence significative du milieu sur les déplacements chimiques. En effet, il aurait été possible que les déplacements chimiques soient perturbés et que cela masque la présence des corrélations caractéristiques des HS, ce qui n'est pas le cas.

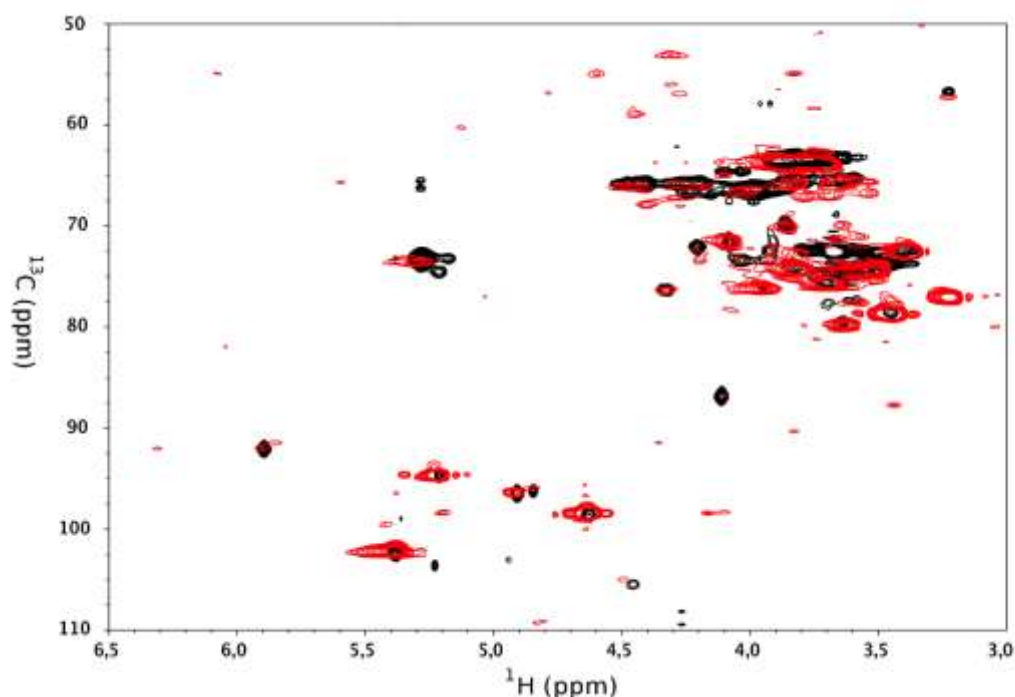


Figure 114 : Comparaison des spectres de membranes HeLa marquées par ssNMR (rouge) et par RMN liquide (noir). Le spectre noir de type HSQC est issu d'une préparation de membranes solubilisées dans l'octylPOE, et le spectre rouge est une analyse INEPT par ssNMR de membranes isolées.

Dans ces conditions de préparation d'échantillon, les HS ne sont toujours pas observés par RMN. Une dernière hypothèse serait que se forment, en présence de détergent, des micelles de taille importante. Les PG, et donc les chaînes d'HS, auraient un temps de corrélation beaucoup plus long, ce qui provoquerait un élargissement des signaux tel que les HS pourraient ne pas être observés. Afin de vérifier cette hypothèse, l'échantillon des membranes solubilisées a été digéré *in situ* aux héparinases et chondroïtinase. Pour cela, l'efficacité de digestion d'héparine commerciale par ces enzymes dans ces conditions de milieu a été vérifiée (résultat non-montré). Le traitement aurait dû alors libérer, si des HS étaient présents, les fragments saccharidiques du protéoglycane. Le spectre HSQC de cette digestion est présenté dans la Figure 115B. La comparaison des spectres de l'échantillon non digéré (Figure 115A) et de l'échantillon digéré ne montre pas l'apparition de nouveaux signaux après digestion. Il n'est pas possible de détecter, même au niveau du bruit, de pic correspondant à la signature spectrale des HS.

En conclusion, il n'a pas été possible d'observer le signal des HS par RMN du solide que ce soit sur les cellules entières ou sur des préparations plus concentrées de fragments de membranes. La raison exacte de cette absence n'est pas connue à ce stade de l'étude. Dans le cas des cellules entières (étudiées en RMN solide dans un rotor de 7mm de diamètre), le fait que le nombre de cellules soit similaire à celui dont sont extraits les échantillons pour la RMN en solution ne permet pas d'invoquer seulement un problème de sensibilité (même en tenant compte de la plus grande sensibilité des cryosondes utilisées en RMN du liquide). Pour les préparations membranaires, on doit se poser la question de l'innocuité du mode de préparation des échantillons sur la conservation des protéoglycanes puisque la présence des HS à la surface des cellules a été précédemment démontrée.

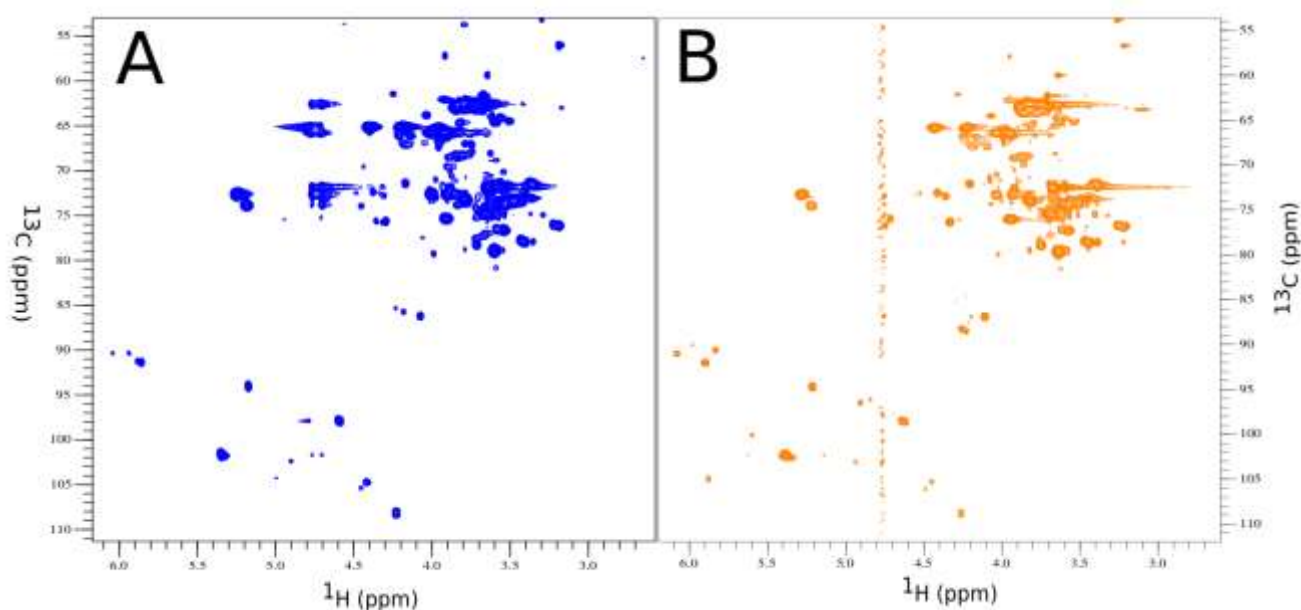


Figure 115 : Analyse par RMN de membranes de cellules HeLa solubilisées à l'octylPOE. Des spectres HSQC sont enregistrés sur les membranes solubilisées à l'octylPOE intactes (A) et digérées aux héparinases et chondroïtinase (B).

Enfin, même si les HS ne sont pas observés, il n'en reste pas moins que des signaux attribuables à des espèces glycosidiques sont présents, et qui sont conservés au niveau des cellules entières et des préparations de membranes solubilisées. La possibilité que ces signaux proviennent de glucose résiduel peut être écartée en raison des lavages successifs des cellules et du mode de préparation des fragments membranaires. Afin de caractériser plus avant les espèces observées dans les spectres HSQC, des expériences préliminaires de type HCCH-COSY et HCCH-TOCSY ont été enregistrées sur un échantillon de membranes solubilisées dans l'octylPOE Figure 116. Ces séquences permettent le transfert d'aimantation soit à travers la première liaison chimique soit à travers toutes les liaisons chimiques dès lors que les noyaux concernés possèdent un spin nucléaire non nul, ce qui est le cas d'un ose entièrement marqué ^{13}C . Cela se traduit dans le spectre par la présence de corrélations à la fréquence d'un proton de corrélations avec le carbone qui le porte et avec son premier voisin (HCCH-COSY) (spectre rouge) ou avec tous les atomes ^{13}C de la chaîne (HCCH-TOCSY) (spectre bleu). Cela est illustré par la Figure 116 où sont indiquées les corrélations de plusieurs motifs différents à partir des

fréquences des protons anomériques. Si l'on considère le système dont le proton est à la fréquence de 5,9ppm (système A), le proton est porté par un carbone à 92ppm, le premier carbone voisin est à 76ppm (on observe une corrélation dans le spectre HCCH-COSY), et la chaîne comporte au moins trois autres carbones à 86ppm, 72ppm et 63,5ppm. Les déplacements chimiques sont proches de ceux compilés pour des riboses mais à ce point, il n'est pas possible d'identifier cette espèce. Dans le cas de l'espèce dont les pics sont les plus intenses dans les spectres de RMN solide (système B, fréquence du proton à 5,4ppm), on obtient pour le carbone associé un déplacement chimique à 102ppm, un premier voisin à 74ppm et 4 autres carbones à 79,5 ppm, 75,5 ppm, 71,7 ppm et 63,5 ppm. Les systèmes C et D dont les corrélations $^1\text{H}_1\text{-}^{13}\text{C}_1$ sont également observées dans les spectres INEPT en solide présentent des schémas voisins. Ces déplacements chimiques sont proches de ceux d'unités de type glucose mais il n'a pas été possible d'aller plus loin dans l'identification de ces composés. Il est néanmoins connu qu'il existe dans les membranes cellulaires un certain nombre de composés glycosylés : protéines membranaires et glycolipides par exemple et il est

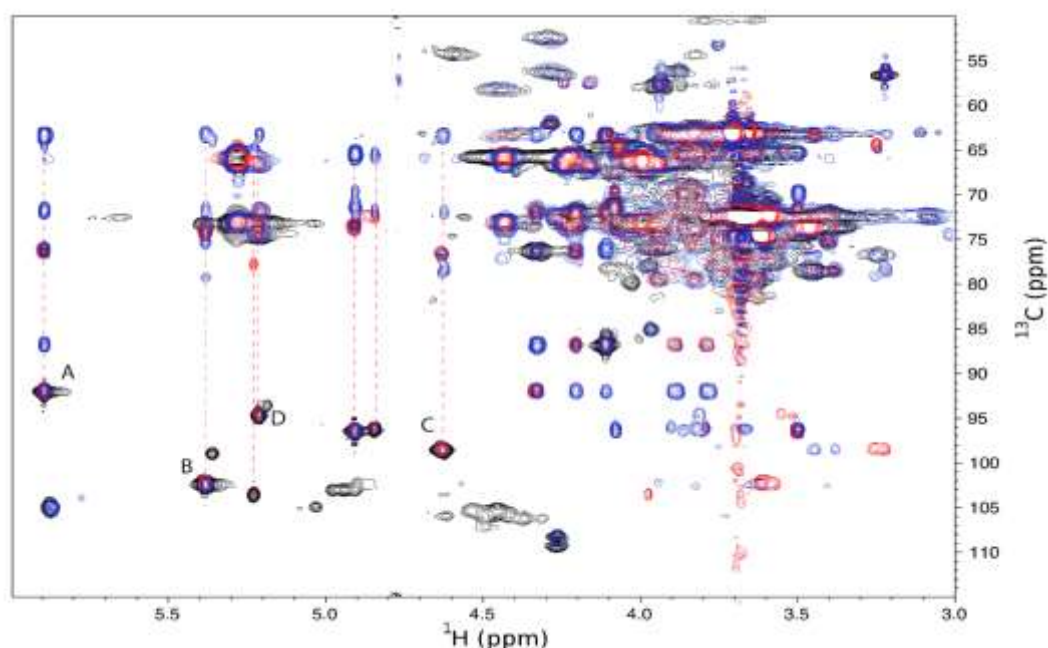


Figure 116 : Analyse par RMN de membranes de cellules HeLa solubilisées à l'octylPOE. Des spectres HSQC (noir), COSY (rouge) et TOCSY (bleu) sont enregistrés sur les membranes isolées solubilisées à l'octylPOE.

possible qu'ils soient à l'origine de ces signaux.

Discussion générale, Conclusion et Perspectives

Les HS sont des acteurs incontournables dans de nombreux événements cellulaires physiologiques et pathologiques tels que la prolifération cellulaire, la migration, le développement, la réponse immune, l'angiogenèse, l'attachement viral *etc* (Sarrazin et al. 2011). En occupant les matrices extracellulaires et les surfaces des cellules, ils permettent de réguler au niveau spatial et temporel la présentation de molécules d'intérêt et de participer à de nombreux processus biologiques. Un grand nombre de protéines est capable d'interagir avec les HS (presque 500 recensées (Ori et al. 2008)), et cette capacité unique des HS prend sa source dans sa structure extrêmement variée. En effet, lors de sa synthèse, et en fonction du type cellulaire (Shi & Zaia 2009), de son activation, de l'état physio-pathologique ou du stade de développement (Bulow & Hobert 2004), le polysaccharide va subir différentes modifications au cours de son élongation et même après au niveau de la surface cellulaire ou dans la MEC. Ces modifications, au niveau du motif disaccharidique (GlcA-GlcNAc), sont principalement des sulfatations et épimérisation à différentes positions, et vont conférer à la chaîne une organisation moléculaire unique avec des domaines fortement substitués alternant avec des régions peu ou pas substituées. Ces domaines hautement modifiés, caractérisés par une très grande hétérogénéité structurale, sont impliqués dans la reconnaissance et l'interaction avec leurs ligands protéiques.

Bien que les HS ne soient pas directement codés génétiquement, la machinerie cellulaire responsable de leur biosynthèse est finement régulée et est dépendante de l'expression d'une large gamme d'enzymes, possédant parfois plusieurs isoformes (Kreuger & Kjellén 2012). Ces enzymes dont l'expression est tissu-dépendante, vont conduire à la formation de polysaccharides de taille, de structure et de dynamique très variables.

Du fait de cette hétérogénéité de structure, de taille, de flexibilité et du manque d'outils analytiques, les analyses structurales des HS ont été fortement limitées. La plupart des études de caractérisation de structure des HS sont principalement basées sur des fragments oligosaccharidiques obtenus par dépolymérisation partielle ou totale de ces polysaccharides. De plus, les analyses disaccharidiques, du fait de la digestion enzymatique ou chimique, font perdre certaines informations structurales, notamment l'épimérisation. Sachant que l'acide iduronique, responsable de la flexibilité de la chaîne d'HS, est extrêmement important pour la reconnaissance des ligands protéiques, il est primordial de développer une approche qui permette également sa caractérisation.

C'est pourquoi, au cours de ma thèse, j'ai développé un protocole pour permettre l'analyse par RMN des HS cellulaires, dans le but de déterminer la structure de chaînes d'HS, non dépolymérisées, extraites de surfaces cellulaires, où elles jouent un rôle majeur.

A partir d'un modèle cellulaire humain, les cellules épithéliales adhérentes HeLa, connues pour exprimer une quantité importante d'HS, nous avons pu mettre au point un nouveau protocole permettant l'analyse structurale par RMN de chaînes d'HS entières marquées. Du fait de la faible abondance naturelle du carbone 13, la stratégie de marquage a été d'utiliser cet isotope sous la forme de ^{13}C -glucose pour marquer métaboliquement les GAGs cellulaires. L'utilisation de l'isotope ^{15}N aurait pu être envisagé, en incorporant de la

glutamine marquée ^{15}N dans le milieu de culture, cependant ce marquage est grandement limité pour la caractérisation des HS (Pomin et al. 2010).

Nous avons pu extraire une quantité d'HS marqués de l'ordre d'une dizaine de microgrammes, après optimisation des conditions de culture et de marquage métabolique au ^{13}C -glucose, et différentes étapes de purification. Cette dernière a reposé principalement sur une série de chromatographies échangeuses d'anions et de digestions des protéines et des acides nucléiques dont la structure linéaire et polyanionique est similaire aux HS. La purification, malgré le fait qu'elle ne soit pas totale, nous permet tout de même de réaliser des études par RMN tant qualitative, par attribution des différents signaux grâce aux données existantes dans la littérature (Guerrini et al. 2005), que quantitative en déterminant le taux de modification ainsi que les structures disaccharidiques. Il a ainsi été mesuré par cette méthode RMN, pour les HS de cellules HeLa, des taux de modification similaires à ceux retrouvés dans la littérature, notamment pour la *N*-sulfatation (Takegawa et al. 2011) utilisant d'autres méthodes, basées principalement sur l'analyse disaccharidique. Cependant, du fait des digestions enzymatiques (héparinases), l'information d'épimérisation n'est plus accessible et nécessite des techniques de dégradation chimique qui elles, font perdre l'information de *N*-sulfatation. Par la technique RMN, ces deux informations (détection des acides iduroniques et des *N*-sulfatations) sont parfaitement accessibles et les HS sont préservés au cours de l'analyse. Les taux de 2-*O*-sulfatation, 6-*O*-sulfatation sont également accessibles ainsi que les différentes unités disaccharidiques formées. La longueur des chaînes d'HS des cellules HeLa a également été estimée entre 70 et 100 disaccharides à partir des intégrations des signaux des glucosamines et du xylose (un xylose par chaîne).

Ce travail de mise au point a permis d'obtenir des profils d'HS cellulaires, grâce à une méthode non destructive, la RMN. Cette technique, du fait de cet avantage important, permet d'utiliser ces HS marqués pour d'autres études, notamment d'interaction avec des ligands protéiques. Un autre avantage majeur de cette méthode est d'avoir accès à des résidus, comme l'acide iduronique, difficilement quantifiable par d'autres techniques.

Pour faire suite à ce travail, il était nécessaire d'utiliser d'autres systèmes cellulaires pour montrer l'efficacité de cette approche. Je me suis donc intéressé à d'autres cellules, les fibroblastes de souris embryonnaires ou MEF, et à l'étude de la structure des HS en fonction de l'expression d'une enzyme impliquée dans la biosynthèse des HS, la NDST-1, qui est vitale *in vivo* (Ringvall et al 2000 et Fan et al 2000). Les analyses *in vitro* par RMN des HS marqués, issus des MEF KO (ou *ndst-1*^{-/-}) comparés aux HS des MEF WT montrent une structure des chaînes différentes bien au-delà de la seule *N*-sulfatation. En effet, dans un premier temps les chaînes d'HS sont effectivement moins *N*-sulfatées (baisse de 36%), cette différence de structure n'a pas pu être observée par cytométrie en flux. Toutefois, du fait de la présence d'autres isoformes de l'enzyme, principalement la NDST-3 et -4 qui sont exprimées plutôt dans les tissus embryonnaires (Aikawa & Esko 1999, Aikawa et al 2001), pourraient être à l'origine des groupements *N*-sulfates observés. Les différents niveaux d'expression ou de surexpression dans les MEF KO peuvent être déterminés pour les différentes isozymes à travers la quantité de transcrits (par PCR quantitative). La NDST-3, comme cela a été décrit dans la partie introduction, a une activité *N*-désacétylase plus importante que son activité *N*-sulfotransférase

(Norgard-Sumnicht & Varki 1995) (Vanpouille et al. 2007), ayant pour conséquence la formation de glucosamines non-substituées (GlcNH_3^+). Toutefois nos analyses par RMN ne nous permettent pas d'observer les déplacements chimiques de ces espèces. L'absence de la NDST-1 et/ou la réduction de la *N*-sulfatation a également des répercussions plus larges sur la structure des HS. En effet, la diminution de la quantité de substrat de la C₅-épipimérase (GlcA-GlcNS) à travers la baisse de la *N*-sulfatation a pour conséquence une baisse de l'épimérisation de 60%, ce qui est plus important que la baisse de *N*-sulfatation. Cela peut s'expliquer par l'élaboration des structures d'HS avec des GlcNS contiguës par la NDST-1, substrat plus favorable pour la C₅-épipimérase. De plus, cette diminution de l'épimérisation ne change pas le niveau de sulfatation des IdoA, qui reste à 65% dans les chaînes d'HS issues des MEF WT et MEF KO. Ceci pourrait suggérer que la HS2ST n'est pas l'enzyme limitante, mais plutôt son association avec la C₅-épipimérase qui est critique. Cette analyse permet également de rendre compte de la spécificité du substrat de la HS2ST qui est l'acide iduronique, indépendamment des modifications des glucosamines adjacentes. La 6-*O*-sulfatation moyenne, quant à elle, reste constante à 20% dans les chaînes d'HS des cellules sauvages et KO, bien que pour ces dernières, la *N*-sulfatation et autres modifications soient réduites. Plusieurs hypothèses peuvent être formulées quant au maintien constant de ce taux de 6-*O*-sulfatation malgré la baisse de la *N*-sulfatation. En effet, cela pourrait s'expliquer par une (sur)expression des isoformes des HS6ST plus affines pour les GlcNAc et surtout pour les GlcNS-IdoA dont la proportion GlcNS,6S-IdoA est nettement supérieure aux GlcNS-IdoA chez les MEF KO (90%) comparé aux MEF WT (35%). L'estimation des différentes modifications, à l'exception de l'épimérisation, réalisée par RMN est en adéquation avec l'analyse disaccharidiques des chaînes d'HS des MEF WT et KO effectuée par nos collaborateurs (Pegeot et al, accepté, voir annexe 7), validant davantage l'efficacité de la purification des HS entiers cellulaires et des méthodes d'analyse par RMN. Aucune évolution de la taille moyenne des chaînes d'HS n'a été observée. L'absence de l'enzyme NDST-1 a un effet sur la cellule qui est important, tout d'abord du point de vue phénotypique et morphologique, les cellules sont beaucoup plus étalées sur leur support, mais le temps de génération est également plus long. L'absence de la NDST-1, en perturbant la machinerie, voire le gadosome (Esko & Selleck 2002), de biosynthèse ordonnée des HS, aboutirait à des structures de chaîne d'HS moins capables d'interagir avec les facteurs nécessaires pour la croissance cellulaire. Ces structures particulières semblent donc être vitales *in vivo* pour la survie du fœtus et du nouveau-né (Ringvall et al. 2000 et Fan et al. 2000). Ceci met également l'accent sur la relation étroite qui existe entre les différentes enzymes de biosynthèse des HS.

Ce protocole de marquage et de purification des HS cellulaires validé par le modèle cellulaire MEF peut également être utilisé pour la détermination des profils de sulfatation et d'épimérisation de différents types cellulaires et au cours de différents phénomènes. En effet, le profil saccharidique des HS en subissant différentes modifications en réponse aux stimulations de l'environnement, permet de moduler les interactions HS-protéines au niveau des surfaces cellulaires. Ces changements dans la voie de biosynthèse et donc dans la structure des HS, qui conduisent à un changement ou perte de fonction de certains ligands protéiques, peuvent se produire lors d'état pathologique ou à travers un phénomène de régulation

cellulaire. Il est donc important de caractériser les chaînes d'HS dans le but de prédire la réponse cellulaire à différents stimuli, notamment en présence de facteurs de croissance et de cytokines avec lesquels elles interagissent et qui sont essentiels dans le développement, la prolifération, la différenciation (Kraushaar et al. 2013) (Casu et al. 2010)...

Le protocole ainsi mis en place a été appliqué pour l'étude de l'évolution du profil d'HS en fonction d'un changement d'état de la cellule, la différenciation cellulaire. Un modèle de cellules Caco-2 issues d'adénocarcinome de côlon a alors été utilisé, ces cellules ayant la capacité de se différencier spontanément en cellules similaires aux entérocytes. Une fois marqués et purifiés, les HS des cellules non-différenciées et différenciées présentent une structure polysaccharidique différente. En effet, les chaînes sont plus courtes et les O-sulfatations et l'épimérisation sont plus importantes dans les HS de cellules différenciées. Le taux de N-sulfatation au cours de la différenciation est stable ce qui est également observé par analyse disaccharidique dans la littérature (Salmivirta et al. 1998). Dans cette étude, le taux de N-sulfatation est toutefois plus faible ($\approx 40\%$) que celui que nous avons déterminé par RMN et que celui obtenu et publié par une autre équipe ($\approx 50\%$) (Levy et al. 1988). Le taux de 6-O-sulfatation dans l'étude de Salmivirta *et al.* est plus faible dans les HS issus de cellules différenciées par rapport aux HS de cellules non-différenciées (-20%) alors que nous observons l'inverse dans notre étude ($+36\%$). De manière générale, à l'exception de la N-sulfatation, les chaînes d'HS des cellules différenciées sont plus modifiées, la proportion de disaccharides GlcNS($\pm 6S$)-IdoA,2S est également plus importante (Pegeot et al. 2014). Ensemble, ces résultats montrent effectivement une évolution de la structure des HS au cours de la différenciation, pour permettre aux cellules d'interagir différemment avec leur environnement. La longueur des chaînes est également réduite d'un facteur 2 approximativement. Toutefois, on ne sait pas si le niveau d'expression en HS est modifié, c'est-à-dire si le raccourcissement observé des chaînes est compensé par l'augmentation du nombre de chaînes. Les analyses par cytométrie en flux pour un même nombre de cellules non-différenciées et différenciées montrent une quantité de fluorescence similaire pour les cellules marquées avec l'anticorps spécifique dirigé contre les HS. La détection des HS par cytométrie en flux ne permet pas une quantification exacte du nombre de chaînes présentes à la surface de la cellule. Ceci est d'autant plus difficile que la structure des chaînes HS est modifiée après différenciation cellulaire.

Un moyen de vérifier si un phénomène de compensation est effectivement présent au niveau des cellules différenciées serait de regarder la variation du taux d'expression des différents PG (glypicans et syndécans principalement) présents sur les surfaces cellulaires. Il serait également intéressant de regarder si les niveaux d'expression des différentes enzymes impliquées dans la biosynthèse des HS varie chez les cellules Caco-2 avant et après différenciation comme c'est le cas pour d'autres enzymes cellulaires (Reisher et al. 1993).

Limites et optimisation de la technique :

Nous avons ainsi pu voir que différentes structures de chaînes d'HS peuvent efficacement être analysées. Toutefois, l'approche utilisée souffre de plusieurs limitations qui peuvent rendre l'analyse des HS de certaines cellules inaccessible. La technique RMN, même si elle a l'avantage d'être non destructive, est peu sensible. En effet, l'analyse requiert une quantité importante d'HS marqués, et donc une grande quantité de cellules, exprimant beaucoup d'HS à leur surface. Bien que la plupart des signaux correspondant aux espèces attendues soient observés, certaines espèces restent encore inaccessibles et non détectées dans les conditions expérimentales déterminées. Ce sont principalement les signaux correspondant aux GlcA,2S, GlcN,3S et GlcNH₃⁺. Ces rares modifications au sein de la chaîne d'HS jouent pourtant un rôle extrêmement important dans l'interaction avec des protéines. Ceci a été plus particulièrement décrit dans le cas de l'ATIII ou encore pour l'HSV (Petitou et al. 2003) (Spear 2004). Si ces résidus sont effectivement présents sur les chaînes d'HS, leurs observations nécessitent une amélioration des conditions expérimentales. Cela peut passer par une augmentation du temps d'expérience, ou l'utilisation de spectromètres à plus haut champ, permettant l'analyse de plus petits volumes, dans des conditions de concentration en HS plus importantes. N'ayant pas eu, au cours de ce travail, la possibilité d'utiliser un tel équipement, nous avons tenté d'augmenter la quantité de matériel. Pour cela, nous avons essayé de mettre au point une méthode de culture des cellules HeLa en condition de non-adhérence afin d'augmenter la quantité de cellules obtenues par unité de volume de milieu marqué, et donc de quantité d'HS. Les premiers essais réalisés n'ont pas permis d'obtenir une lignée viable et nous sommes donc restés dans des conditions d'adhérence. Par conséquent, du fait de ce manque de sensibilité, il n'est pas encore possible de détecter les espèces rares, et il est très difficile d'appliquer un tel protocole et de réaliser ce genre d'étude sur des cellules exprimant très peu d'HS comme par exemple, les lignées lymphocytaires.

De plus, cette limite de sensibilité est encore accentuée par un marquage non spécifique, à savoir l'utilisation du précurseur ¹³C-glucose pour le marquage des HS. L'inconvénient majeur de ce type de marquage est d'introduire des isotopes ¹³C du carbone au niveau des glycoconjugués et des acides nucléiques, nécessitant plusieurs étapes de purification. Bien que les échantillons d'HS ainsi préparés permettent des analyses qualitatives et quantitatives, des expériences d'interaction en RMN devraient requérir toutefois une purification plus fine des extraits. Celle-ci se basant principalement sur l'élimination des acides nucléiques restants dont les déplacements chimiques de leur désoxyribose/ribose ne sont pas suffisamment résolus et se confondent avec les déplacements chimiques des HS, surtout au niveau des corrélations ¹H₂₋₅-¹³C₂₋₅ (*Biological Magnetic Resonance Data Bank*). De plus, du fait que HS et ADN/ARN soient des polymères anioniques, les interactions de natures électrostatiques HS-protéines peuvent être parasitées en présence d'acides nucléiques. Les chaînes d'HS purifiées sont également toujours porteuses d'une fraction peptidique de la protéine cœur au niveau du site d'ancrage serylglycine. Cette fraction peut être détachée du xylose par voie chimique en utilisant la β-élimination ménagée. Enfin, une dernière étape de purification pour l'obtention d'HS cellulaires ultra purs seraient de digérer les CS présents

(moins de 10%) dans l'échantillon et observés en RMN mais pas par cytométrie en flux.

La sensibilité du signal des HS par RMN étant cependant suffisante pour déterminer l'essentiel du profil saccharidique, des analyses des HS par RMN du solide ont été mises en place. Les analyses des cellules HeLa entières et des surfaces membranaires ont clairement montré la présence de sucres relativement flexibles au niveau des membranes dans les différentes conditions. Ces sucres marqués, du fait de leur déplacement chimique ne semblent pas appartenir aux HS. La méthode montre cependant qu'il est possible de sonder une partie de la surface cellulaire par ssNMR après marquage selon le protocole présenté. Dans le but d'augmenter la sensibilité des expériences ssNMR, des essais de polarisation nucléaire dynamique (DNP) ont également été mis en place d'après le protocole utilisé pour les cellules procaryotes (Takahashi et al. 2013). En dehors de toute optimisation du protocole pour les cellules HeLa, le spectre montre différentes corrélations dont deux pourraient être dues à la présence effective de glucosamine et d'acide glucuronique. Toutefois, afin de clairement déterminer leur présence, des optimisations sont nécessaires du point de vue de la préparation des échantillons et du marquage ainsi que des améliorations spectroscopiques au niveau des séquences d'impulsion par exemple ou d'analyser les cellules et leur surface dans des spectromètres à plus haut champ. Dans le cas de la RMN du solide sur des échantillons de membranes cellulaires, pour améliorer la sensibilité, une piste serait de reconsidérer le marquage indirect, et d'utiliser des précurseurs d'HS plus spécifiques pour s'affranchir de tous ces contaminants et augmenter l'efficacité de marquage.

Ensemble, les différents résultats obtenus au cours de ma thèse montrent qu'il est possible d'étudier les structures des HS après un marquage métabolique pour leurs analyses par RMN liquide. Ce protocole de marquage peut être appliqué à différentes lignées cellulaires et au cours de différents événements cellulaires. Les cellules doivent cependant exprimer une certaine quantité d'HS à leur surface pour pouvoir les analyser par RMN en phase liquide. Il est donc possible, techniquement de suivre l'évolution structurale des GAG suite à un changement d'état de la cellule (pathologique ou régulation physiologique). Cette approche pourrait être particulièrement intéressante dans le cas de cellules souches ou cellules cancéreuses, pour lesquelles la croissance et la différenciation cellulaire sont sous le contrôle d'un éventail de ligands protéiques fixant les HS, comme les facteurs de croissance, cytokines et morphogènes (Casu et al, 2010).

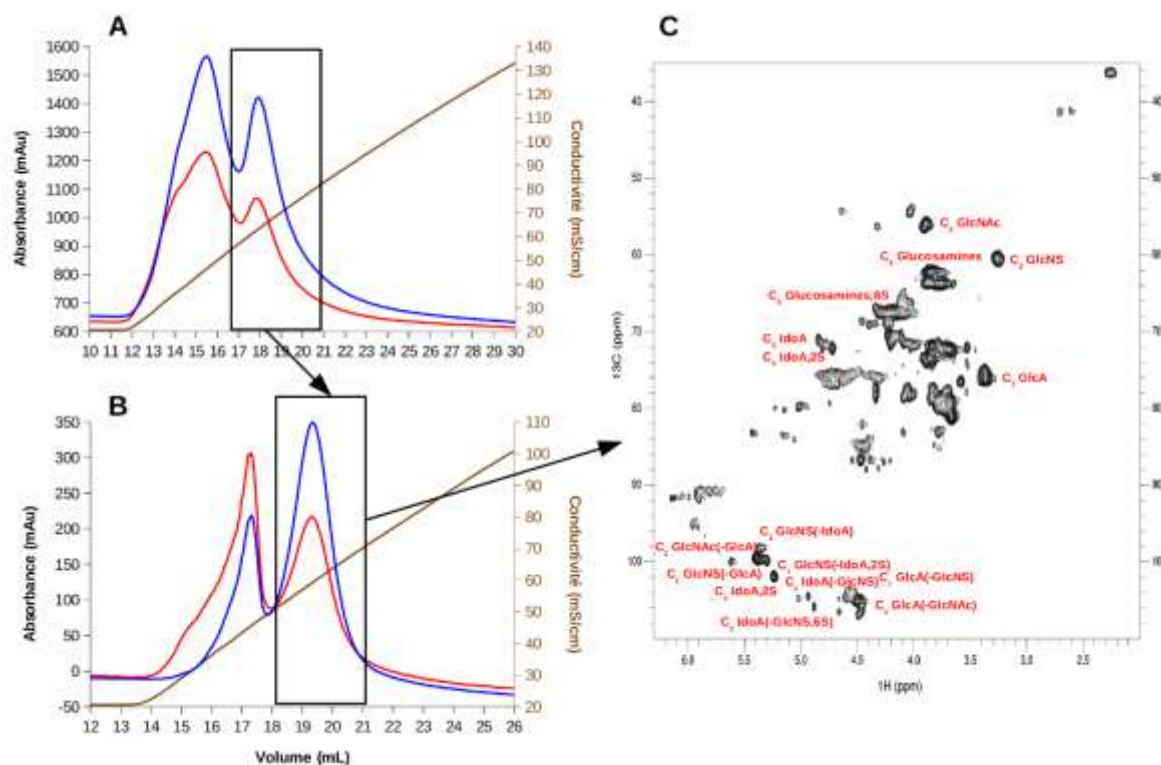
Un autre avantage de cette technique qui permet de générer des HS marqués au ^{13}C est leur potentielle utilisation pour des études d'interaction avec un ligand protéique afin d'identifier les déterminants structuraux impliqués dans cette association.

L'investigation des structures d'HS par RMN du solide reste une voie d'exploration qui ne permet pas à l'heure actuelle d'observer les HS directement sur des cellules ou membranes cellulaires isolées. Des techniques de DNP et un marquage plus spécifique pourraient améliorer la sensibilité des expériences pour permettre la détection effective des HS dans le contexte membranaire.

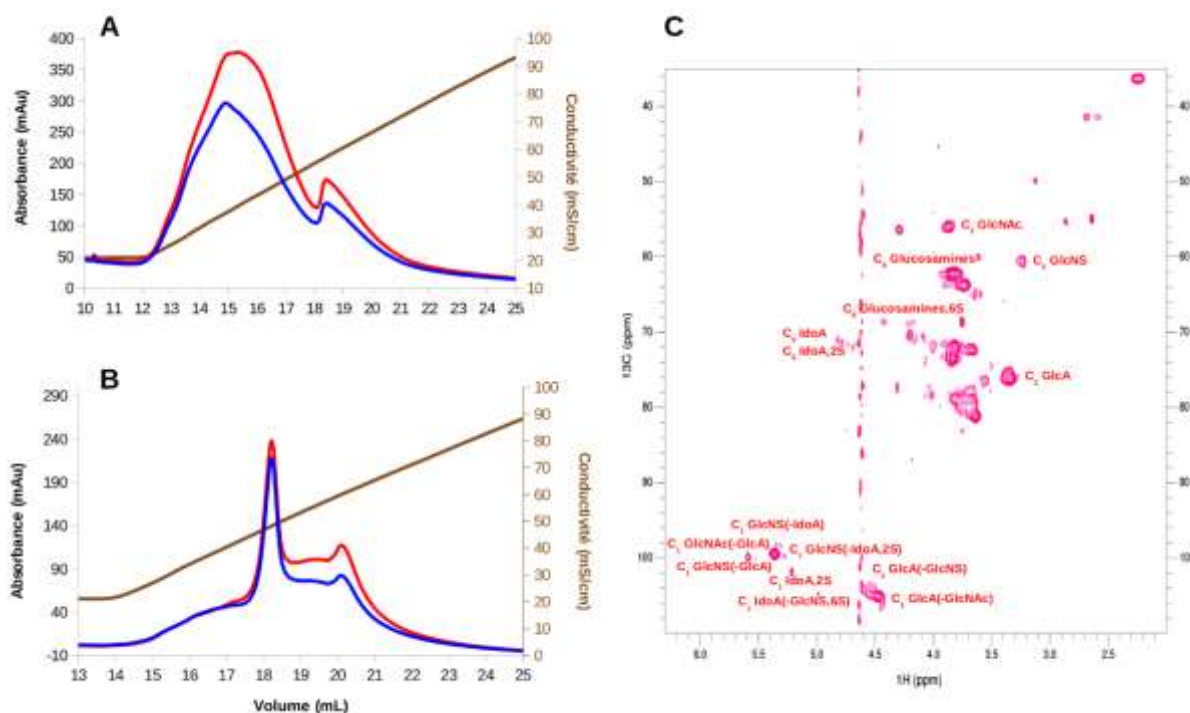
Annexes

Topologie	Protéoglycane	PM (hors glycosylation)	Type de OAG	Nombre de chaînes	Localisation(s) principale(s)
MEC	Aggrecan	250kDa	CS, KS I KS II	~100(CS), 2-3(KS I), ~20(KS II)	Cartilages, neurones
	Biglycan	42kDa	DS	2	Tissus conjonctifs, cartilages articulaires
	Collagène IX	92kDa	CS/DS	1	Cartilage hyalin
	Collagène XII	33kDa	CS	3 sites	Muscles squelettiques, intestin grêle
	Collagène XV	142kDa	CS(HS)	11 sites	Pancréas, reins
	Collagène XVIII	178kDa	HS	1-3	Cellules épithéliales, membranes basales
	Décronine	48kDa	CS ou DS	1	Tissus conjonctifs
	Epiphygan	37kDa	CS, DS	2-3	Cartilages, ligaments, placenta
	Fibromoduline	43kDa	KS I	4	Cartilages
	Keratocan	41kDa	KS I	3-5	Cornée
	Lubricine	151kDa	CS	1	Liquide synovial, articulations
	Lumican	38kDa	KS I	3-4	Cornée
	Neurocan/CSPG3	143kDa	CS	3-7	Cerveau
	Ostéoglycine/Minécan	34kDa	KS I	2-3	Cornée, os
	Ostéomoduline/Ostéoadhérine	49kDa	KS I	2-3	Os
	Perlecan/HSPG2	469kDa	HS	3	Membranes basales, cartilages
	Phosphacan/PTPRZ1	164kDa/258kDa	CS	4	Système nerveux central (pour les 2 isoformes)
	PRELP/Prolagin	44kDa	KS I	2-3	Tissus conjonctifs, cartilages
	Testican-1	48kDa	HS, CS	2	Cerveau, appareil reproducteur mâle
	Testican-2	47kDa	HS, CS	2	Cerveau, poumons, testicules
	Testican-3	49kDa	HS	2	Cerveau
	Versican (V0-V2, V4)	115-373kDa	CS, DS	12-23	Fibroblastes, tissus conjonctifs
MEC/Membrane Plasmique	Agrin	217kDa	HS	2-3	Membranes basales des poumons et reins, neurones
	Bilaglycan	93kDa	HS(CS)	1	Expression tissulaire large
	Brevican/CSPG7	99kDa ou 72kDa	CS	3	Cerveau (cellules gliales, neurones)
Membrane Plasmique	Glypican-1 (2 isoformes)	62kDa/32kDa	HS	3(1)	Fibroblastes, cellules épithéliales
	APLP2 (Amyloid precursor like protein 2)	87kDa	HS/CS	1	Placenta, cerveau, endothélie
	Appican (11 isoformes)	87kDa	CS	1	Tissus fœtaux, cerveau adulte
	CD44 (19 isoformes)	82kDa (CD44s)	CS	1	CD44s : cellules hématopoïétiques, leucocytes, fibroblastes
	CD44v3	37kDa	HS	1	Lymphocytes
	Cérébroglycan/Glypican-2	63kDa	HS	5	Système nerveux en développement
	Fibroglycan/Syndécan-2	22kDa	HS	2-3	Expression tissulaire large
	Glypican-5	64kDa	HS	4	Cerveau
	Glypican-6	63kDa	HS	4	Large expression
	K-Glypican/Glypican-4	62kDa	HS	3	Système nerveux central, reins
	Neuroglycan C/CSPG5	68kDa	CS	1	Cerveau
	NO2/CSPG4	250kDa	CS	1	Cellules de mélanomes
	N-Syndécan/Syndécan-3	48kDa	HS/CS	3-4/1-2	Système nerveux, rate, glandes surrénales
	OCI-5/MOR1/Glypican-3	66kDa	HS	2	Poumons, foie, reins
	Ryudocan/Amphiglycan/Syndécan-4	22kDa	HS	2-3	Cellules épithéliales et fibroblastes
	Syndécan-1	32kDa	HS/CS	3-4/1-2	Cellules épithéliales
	Thrombospondine	60kDa	CS, DS	1(CS), 1(DS)	Cellules endothéliales
Intracellulaire	Serglycine	18kDa	CS/Hp	8	Cellules hématopoïétiques, endothéliales, tissus conjonctifs
Circulation	ApoD	22kDa	CS	1	Sang, ubiquitaire
	Bikunine	18kDa	CS	1	Sang
	Endocan/ESM-1	20kDa	DS	1	Sang, endothélie
	PG-109/MCSF1	60kDa	CS	1	Sang
Noyau/MEC	Bamacan/CSPG6	142kDa	CS	3	Membranes basales

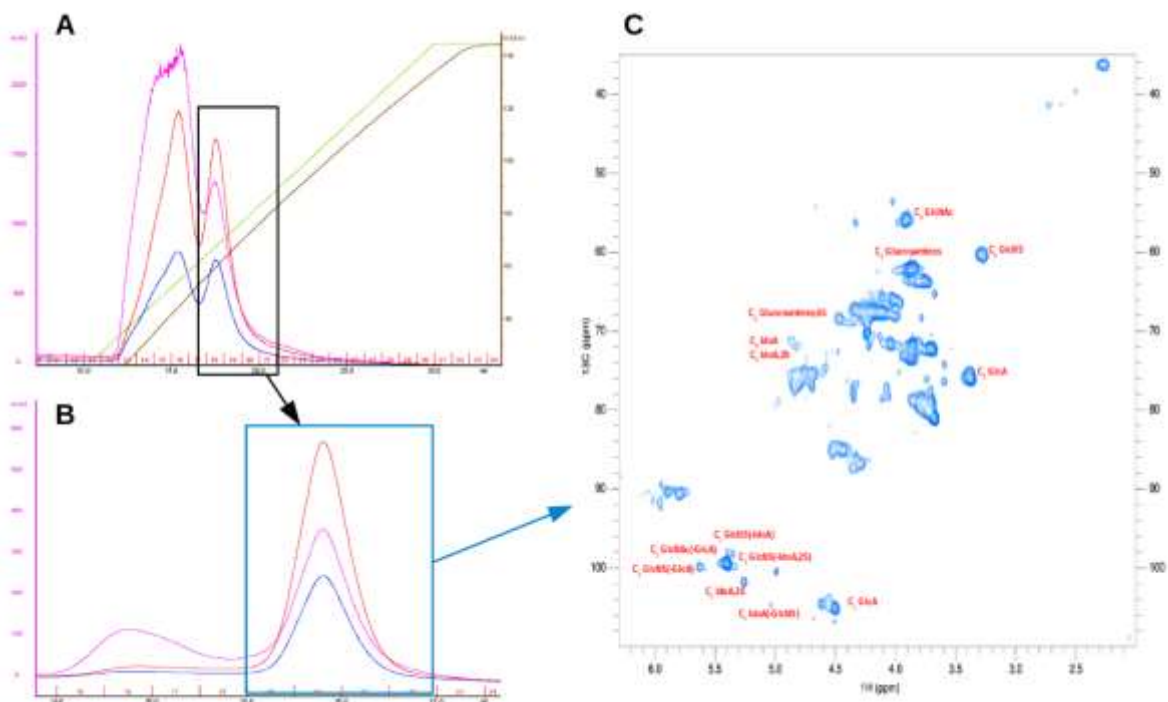
Annexe 1 : Liste non-exhaustive des différents PG présents chez l'Homme. Les PG sont classés selon leur topologie dans l'organisme. Les PG sont ensuite classés par ordre alphabétique. Les poids moléculaires (PM) hors glycosylation, le type et le nombre de chaîne(s) porté(s) ainsi que les tissus d'expression principaux sont également renseignés.



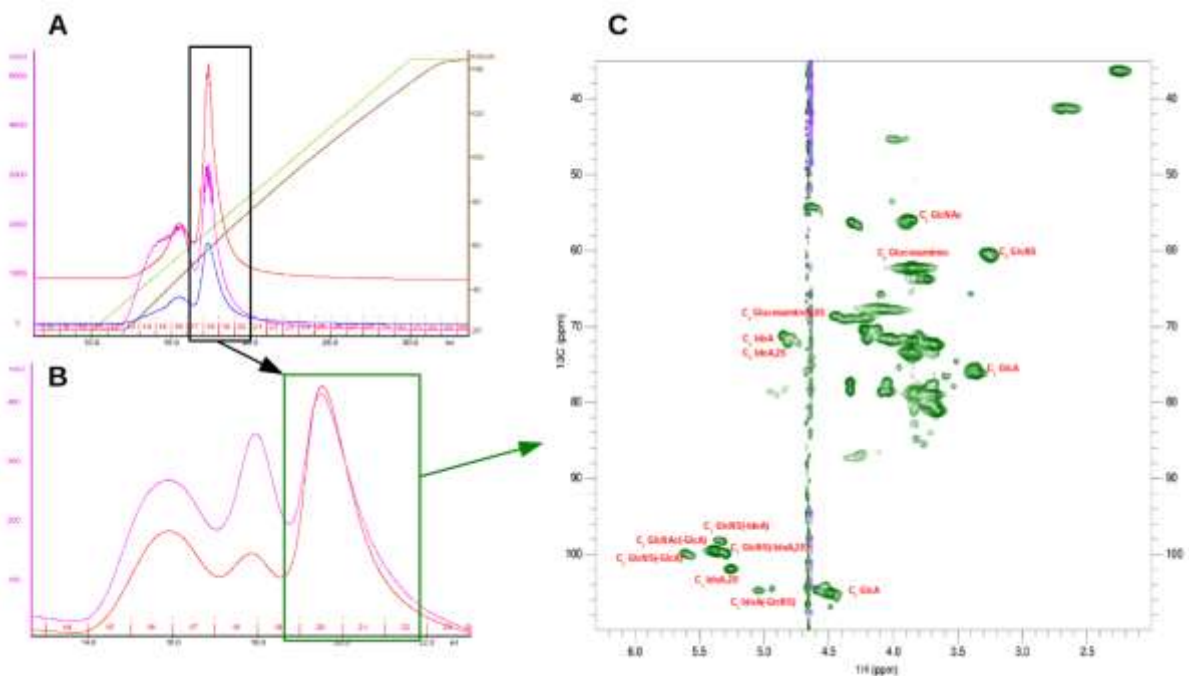
Annexe 2 : Purification et analyse RMN d'un échantillon-type d'HS issus de MEF WT. Les extraits d'HS de MEF WT sont purifiés par chromatographies successives après digestion à la trypsine (A) et à la DNase I (B). Seules les fractions éluées en forte concentration de sel absorbant à 260nm (rouge) et 280nm (bleu) sont gardées pour analyse par RMN. Les spectres sont enregistrés sur un spectromètre 700MHz à 37°C pendant 16h et attribués d'après Guerrini et al 2005 (C)



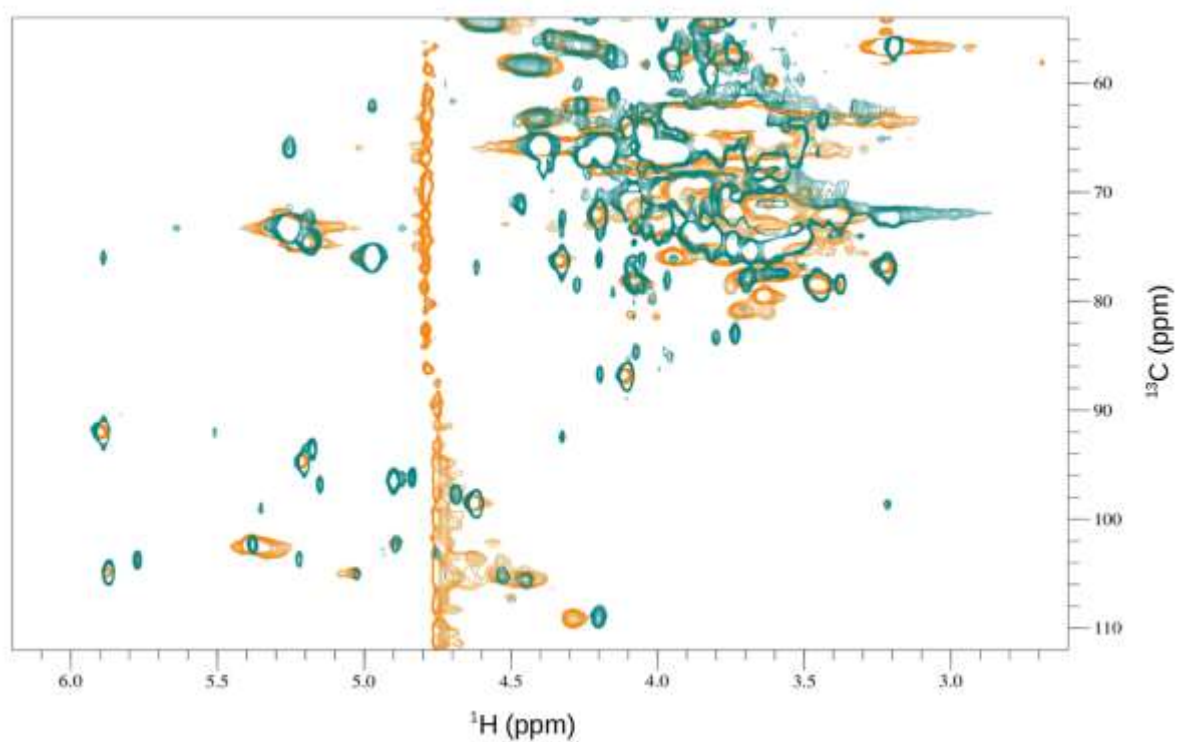
Annexe 3 : Purification et analyse RMN d'un échantillon-type d'HS issus de MEF ndst-1^{-/-}. Les extraits d'HS de MEF KO sont purifiés par chromatographies successives après digestion à la trypsine (A) et à la DNase I (B). Toutes les fractions d'éluion absorbant à 260nm (rouge) et 280nm (bleu) sont gardées pour analyse par RMN. Les spectres sont enregistrés sur un spectromètre 600MHz à 37°C pendant 16h et attribués d'après Guerrini et al 2005 (C)



Annexe 4 : Purification et analyse RMN d'un échantillon-type d'HS issus de Caco-2 non-différenciées. Les extraits d'HS de Caco-2 non-différenciées sont purifiés par chromatographies successives après digestion à la trypsine (A) et à la DNase I (B). Seules les fractions éluées en forte concentration de sel absorbant à 260nm (rouge), 280nm (bleu) et 215nm (magenta) sont gardées pour analyse par RMN. Les spectres sont enregistrés sur un spectromètre 700MHz à 37°C pendant 15h et attribués d'après Guerrini et al 2005 (C)



Annexe 5 : Purification et analyse RMN d'un échantillon-type d'HS issus de Caco-2 différenciées. Les extraits d'HS de Caco-2 différenciées sont purifiés par chromatographies successives après digestion à la trypsine (A) et à la DNase I (B). Seules les fractions éluées en forte concentration de sel absorbant à 260nm (rouge), 280nm (bleu) et 215nm (magenta) sont gardées pour analyse par RMN. Les spectres sont enregistrés sur un spectromètre 600MHz à 37°C pendant 15h et attribués d'après Guerrini et al 2005 (C)



Annexe 6 : Spectres ^1H - ^{13}C HSQC de membranes isolées solubilisées en détergents. Les membranes isolées de cellules HeLa sont analysées par RMN liquide après solubilisation dans du Triton X-100 (vert) ou DPC (orange). Les spectres sont enregistrés à 25°C sur un spectromètre 600MHz équipé d'une crysonde pendant 3h.



**Profiling sulfation/epimerization pattern of full-length
heparan sulfate by NMR following cell culture ¹³C-glucose
metabolic labeling**

Journal:	<i>Glycobiology</i>
Manuscript ID:	GLYCO-2014-00173.R1
Manuscript Type:	Communications
Date Submitted by the Author:	15-Oct-2014
Complete List of Authors:	Pegeot, Mathieu; Institut de Biologie Structurale, SAGAG Sadir, Rabia; Institut de Biologie Structurale, SAGAG Eriksson, Inger; Uppsala University, Medical biochemistry and microbiology Kjellen, Lena; Uppsala University, Medical biochemistry and microbiology Gans, Pierre; Institut de Biologie Structurale, SAGAG Simorre, Jean-Pierre; Institut de Biologie Structurale, SAGAG Lortat-Jacob, Hugues; Institut de Biologie Structurale, SAGAG
Key Words:	Heparan sulfate, Sulfation pattern, GlcA/IdoA ratio, ¹³ C-metabolic labelling, NMR

SCHOLARONE™
Manuscripts

**Profiling sulfation/epimerization pattern of full-length heparan sulfate by NMR
following cell culture ^{13}C -glucose metabolic labeling**

Mathieu Pegeot^{1,2,3}, Rabia Sadi^{1,2,3}, Inger Eriksson⁴, Lena Kjellen⁴, Jean-Pierre Simorre^{1,2,3},
Pierre Gans^{1,2,3,*}, Hugues Lortat-Jacob^{1,2,3,*}

1 Univ. Grenoble Alpes, Institut de Biologie Structurale (IBS), F-38027 Grenoble, France

2 CNRS, IBS, F-38027 Grenoble, France

3 CEA, DSV, IBS, F-38027 Grenoble, France

4 Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala,
Sweden.

Corresponding Authors:

*Pierre Gans (Pierre.Gans@ibs.fr) and *Hugues Lortat-Jacob (Hugues.Lortat-Jacob@ibs.fr;
Tel: +33 457 428 533; Fax: +33 476 501 890).

Running head: NMR profiling of cellular HS sulfation/epimerization pattern

Key words: Heparan sulfate/Sulfation pattern/Quantification of GlcA/IdoA
epimerization/ ^{13}C -metabolic labelling/NMR

Supplementary data (supplementary figures 1 to 5) for this article are available.

ABSTRACT

Through its ability to interact with proteins, heparan sulfate (HS) fulfills a large variety of functions. Protein binding depends on the level of HS sulfation and epimerization which are cell specific and dynamically regulated. Characterization of this molecule, however, has been restricted to oligosaccharide fragments available in large amount for structural investigation or to sulfate distribution through compositional analysis. Here we developed a ^1H - ^{13}C two-dimensional NMR based approach, directly performed on HS isolated from ^{13}C -labeled cell. By integrating the peak volumes measured at different chemical shifts, this non-destructive analysis enables to determine both the sulfation and the iduronic/glucuronic profiles of the polysaccharide. Applied to wild type and *N*-deacetylase/*N*-sulfotransferase-deficient fibroblasts as well as to epithelial cells differentiation, it also gives insights into the functional relationships existing between HS biosynthetic enzymes. This approach should be of large interest to better understand HS changes that occur through physiologic regulations or during pathological development.

INTRODUCTION

Heparan sulfate (HS) is a complex polysaccharide of the glycosaminoglycan (GAG) family, that occurs covalently attached to core proteins both at the cell surface and within extracellular matrices. Through its ability to bind to a large array of proteins, almost 500 of which have been identified to date (Ori, A., Wilkinson, M.C., et al. 2008), this molecule regulates key biological processes at the cell-tissue interface, such as cell proliferation migration and development, chemoattraction, inflammation and immune response, lipid metabolism, angiogenesis, matrix assembly or viral attachment (Sarrazin, S., Lamanna, W.C., et al. 2011). To do so, HS can either or both provide scaffolds to ensure that proteins mediating specific functions will be presented at the correct site and time, and directly contribute to biological activities or signaling processes. It has been well known, for example, that HS mediates the formation of chemokine gradients along which cells can migrate directionally, provides templates onto which two adjacent proteins such as growth factors and receptors can interact to facilitate signal transduction, protects cytokines against proteolysis, providing reservoirs of active factors that can be mobilized in specific conditions, or induces protein conformational changes thereby modulating their biological functions (Lindahl, U. and Li, J.P. 2009, Sarrazin, S., Lamanna, W.C., et al. 2011).

Consistently with its multiple binding activities, HS features extended structural variability. It typically comprises 20 to 200 alternating glucuronic acid (GlcA) and N-acetyl-glucosamine (GlcNAc) linked residues that, during biosynthesis, are variously modified by interdependent enzymatic reactions. These modifications include GlcNAc *N*-deacetylation and *N*-sulfation, C5 epimerization of GlcA into iduronic acid (IdoA) and various *O*-sulfations, occurring frequently at the C6 and C2 positions of the GlcN and IdoA respectively and more rarely at

the C3 and C2 position of the GlcN and GlcA respectively. A large body of evidence supports that protein binding properties are largely dependent upon HS domains of distinctive sulfation and epimerization pattern (Lindahl, U. and Li, J.P. 2009). HS modifications appeared to be cell, tissue and development stage specific (Bulow, H.E. and Hobert, O. 2006, Ledin, J., Staatz, W., et al. 2004, Shi, X. and Zaia, J. 2009, van Kuppevelt, T.H., Dennissen, M.A., et al. 1998) and are dynamically regulated (Kreuger, J. and Kjellen, L. 2012). Thus, changes in HS biosynthetic pathway and structure that occur through physiologic regulation or during disease state can alter its protein binding capacity, resulting in change or loss of function.

Approaches for the analysis of cellular HS are commonly based on partial or complete depolymerization to produce a range of oligosaccharide fragments amenable to separation and structure elucidation methods such as mass spectrometry (Zaia, J. 2013). Disaccharide compositional profiling of HS, usually radioactively or fluorescently labeled, allows for quantitative analysis by using appropriate disaccharide standards but require significant sample preparation and can give rise to biased results as certain regions of HS have been shown to be resistant to digestion down to disaccharides. In addition, enzymatic depolymerisation causes a loss of the chiral center that distinguishes IdoA from GlcA, thus, the epimerization profile of the starting material cannot be evaluated (Desai, U.R., Wang, H.M., et al. 1993, Galeotti, F. and Volpi, N. 2013). IdoA residues strongly influence protein recognition (Casu, B. and Lindahl, U. 2001), their quantification is therefore of large importance. NMR has also been proved to be a powerful approach to investigate GAG composition. However, due to the weak natural abundance of the ^{13}C isotope (1.1%), this analysis was restricted to polysaccharides available in mg amount such as tissue extracted material (Guerrini, M., Naggi, A., et al. 2005) or to well defined oligosaccharides generated by synthetic chemistry- (Langeslay, D.J., Beni, S., et al. 2012) or chemo-enzymatic- based

(Zhang, Z., McCallum, S.A., et al. 2008) strategies. Here, to overcome these limitations, we used ^{13}C -glucose as a precursor of both GlcN and GlcA to metabolically label GAGs produced by growing cells in culture. With this labeling strategy we established that two-dimensional NMR spectroscopy permits to characterize both the sulfation and epimerization pattern of non-depolymerized HS, directly isolated from cells. This method only requires minimum sample preparation. In particular, as glucosamines and galactosamines give distinct signals in NMR (Huckerby, T.N., Nieduszynski, I.A., et al. 2005), chondroitine sulfate (CS) does not need to be eliminated prior to analysis. Using different cellular models we also showed that this approach can reveal HS modifications that result from changes in HS biosynthetic enzyme level of expression or that occur during cellular differentiation.

RESULTS

^{13}C -glucose metabolic labeling yield HS samples that can be characterized by 2D-NMR

To set up the experimental details, and demonstrate that quantitative HS analysis was possible with a reasonable amount of material, GAGs were first extracted from $\sim 2 \times 10^8$ HeLa cells, a number that can be easily obtained in cell culture. The presence of HS at the cell surface and the efficiency of the extraction were demonstrated by FACS scan analysis (supplementary Figure 1) and solubilized HS were purified by DEAE chromatography. Eluted material resolved in two main peaks, exhibiting oligonucleotide and typical GAG signals respectively (supplementary Figure 2). DNase treatment of this material gave rise to a sample which NMR spectrum exhibited all the correlations expected for GAGs and was assigned in accordance with that recorded on tissue extracted HS (Guerrini, M., Naggi, A., et al. 2005) at the natural ^{13}C abundance (Figure 1a). Comparison with a spectrum of 2 mg of commercial

HS (supplementary Figure 3) enabled us to estimate that ~ 10 µg of ^{13}C -labeled HS were obtained from 2×10^8 cells. Having established that ^{13}C -glucose based metabolic labeling permits recording signals with good resolution from a limited amount of material, we integrated the peak volumes measured at different chemical shifts for the different mono- and di- saccharide units (Figure 1a). Sulfation levels and IdoA/GlcA ratio were calculated from the intensities of the closest NMR probes. Thus, to quantify the *N*-sulfation level, the $\text{H}_2\text{-C}_2$ resolved signals of GlcNS and GlcNAc were used which returned a value of 46%. Comparison of the C6 signals from GlcN6S and GlcN6H indicated that 23% of the GlcN were 6-*O*-sulfated. Finally, the GlcA and IdoA residues giving rise to distinguishable chemical shifts, a direct evaluation of the IdoA/GlcA ratio was possible. The $\text{H}_1\text{-C}_1$ signals from IdoA2S, IdoA-GlcNS and IdoA-GlcNS6S and that from GlcA-GlcNAc and GlcA-GlcNS were quantified and showed that the IdoA/GlcA ratio was 29% with 65% of the IdoA residues being 2-*O*-sulfated (Figure 1b). Overall, the Hela HS chains feature an average of 0.88 sulfate per disaccharide, which is in the typical range (0.57–1.25) of HS (Gallagher, J.T. and Walker, A. 1985). Analysis of the anomeric proton-carbon cross-peaks, which are sensitive to the neighboring monosaccharide, also provides additional information about residues linkages within disaccharide sequences and reveals the proportion of the different residues adjacent to GlcNS, IdoA and GlcA (Figure 1c).

N-deacetylase/N-sulfotransferase silencing has multiple effects on HS structure

Next, we investigated the HS characteristics of WT and N-deacetylase/N-sulfotransferase-1^{-/-} (NDST1^{-/-}) KO mouse fibroblasts. NDST1 is one of the four NDST that catalyze the de-*N*-acetylation/re-*N*-sulfation of GlcNAc residues. The HS synthesized by these cells can be predicted to feature a lower level of GlcNS compared to the WT cells and were thus extracted and analyzed as described above. The HS chains showed clear differences in their sulfation

1
2
3 pattern and disaccharide composition (Figure 2) with, as expected, a significant decrease of
4
5 N-sulfation (51 to 34%) in the KO. A shift in epimerization (from 36 to 14% of IdoA) was
6
7 also noticed. This result is consistent with the biosynthetic process as the C₅-epimerase
8
9 required a GlcNS upstream of its substrate to convert GlcA into IdoA (Hagner-Mcwhirter, A.,
10
11 Lindahl, U., et al. 2000). It can be also calculated that in the WT cells, 73% of the GlcNS
12
13 were linked to an IdoA residue, versus 48% in the KO. This could suggest a sequential
14
15 mechanism of the C₅-epimerase that would be more active when its substrate (GlcA) is
16
17 embedded within stretches of GlcNS than on isolated GlcNS-GlcA sequences. Finally, the
18
19 total 2-*O*-sulfation level was reduced from 24 to 9 % of the hexuronic acids which is
20
21 consistent with the fact that 2-*O*-sulfotransferase preferentially uses IdoA over GlcA as a
22
23 substrate. Nevertheless, IdoA2S still represents 66% of the IdoA residues in both cells,
24
25 indicating that epimerized residues have been sulfated with the same efficiency in both cases
26
27 (Figure 2a). Overall, the HS from WT- and NDST1^{-/-} cells feature an average of 0.95 and 0.61
28
29 sulfate per disaccharide respectively. To support these data and the methodological approach,
30
31 the HS-derived disaccharides from WT- and KO- MEFs were also characterized by reversed
32
33 phase ion-pair HPLC. The result indicated shifts in the *N*- and 2-*O*-sulfation level and no
34
35 change in 6-*O*-sulfation, in perfect agreement with the NMR data (supplementary Figure 4).
36
37 In addition to the IdoA/GlcA ratio, that disaccharide analysis cannot quantify, our approach
38
39 also revealed major differences between these two cell lines. In the WT MEFs, 29, 31 and
40
41 40% of the GlcNS(±6S) were associated to GlcA, IdoA and IdoA2S respectively, whereas in
42
43 the KO cells the corresponding values were 52, 25 and 23%. More importantly, we found that
44
45 65 and 35% of the IdoA residues were followed by a GlcNS and GlcNS6S respectively in the
46
47 WT- versus 12 and 88% in HS from the KO- MEFs (Figure 2b).
48
49
50
51
52
53
54
55
56
57
58
59
60

Caco-2 differentiation induced HS modifications can be followed by NMR

To investigate whether this method can be used to follow HS regulatory processes, we took advantage of the human Caco-2 colon carcinoma cells which undergo spontaneous differentiation when maintained in culture for more than 12 days post-confluency and is frequently used to study intestinal epithelium development and function (Sambuy, Y., De Angelis, I., et al. 2005). NMR spectra were obtained for undifferentiated as well as for differentiated cells. The HS synthesized by non-differentiated Caco-2 cells were very similar to those found on HeLa cells, with 48% of GlcNS, an IdoA/GlcA ratio of 27%, and 16 and 24% of 2-*O* and 6-*O* sulfation respectively, amounting to an average of 0.88 sulfate per disaccharide (Figure 3a). Once differentiated, an increase in GlcA epimerization (27 to 40 %) that was accompanied by an increase in 2-*O* (16 to 26%) and 6-*O* (24 to 31 %) sulfation was observed, giving rise to an average of 1.1 sulfate per disaccharide. More importantly a modification in disaccharide organization was noticed, with GlcNS(\pm 6S) being more frequently linked to IdoA (0.27) and IdoA2S (0.4) in differentiated cells than in non-differentiated ones (0.26 and 0.21 for IdoA and IdoA2S respectively).

HS chain length estimation

Finally, in some cases, an estimation of the chain lengths can also be provided by quantifying the xylose (Xyl) residue of the tetrasaccharide motif linking HS to core proteins. The signal intensity of the Xyl H₃-C₅ proton, which is in a resolved region of the spectrum, was used as an internal reference and compared to the sum of the intensities of the H₂-C₂ signals of GlcNAc and GlcNS residues. For WT and KO MEFs, an equivalent length of the chains with a mean value of about 30 \pm 5 disaccharide units (supplementary Figure 5) was found.

Differentiated Caco-2 cell had the same chain length whereas in undifferentiated Caco-2 and HeLa cells, the Xyl signals although observable were very weak, indicative of longer chains

roughly estimated to contain 70-100 disaccharide units. Regarding the use of this signal for chain length estimation, it should be noted that it can originate from both HS and CS, thus the estimation will be valid only if CS is present as a negligible entity (as demonstrated here; supplementary Figure 1).

For Peer Review

DISCUSSION

Altogether, these results showed that the main HS characteristics, including sulfation/epimerization pattern and chain length can be profiled using NMR based non destructive analysis, directly from materials extracted from cultured cells. Yet, this approach has some limitations: 1) the number of cells required to perform such analysis might be high for those expressing only low amounts of HS such as the Jurkat T cell line for example (data not shown) and sensitivity can also be an issue to detect rare modifications. The present work showed however that the Xyl residue, which occurs only once per chain, can be measured; 2) the use of glucose as a labeled precursor is likely to introduce ^{13}C in a variety of glycoconjugates, preventing a direct HS analysis without a minimum of sample preparation. ^{15}N -isotopic labeled glutamine incorporation has been reported already, allowing the characterization of the sulfamate groups microenvironment in HS extracted from HEK-293T cells (Langeslay, D.J., Beecher, C.N., et al. 2013). Such labeling also permits observation of signals characteristic of GlcNAc in CHO-derived HS as well as that from 4-*O*- and 6-*O*-sulfated N-acetyl-galactosamine in chondroitin sulfate from mouse lung endothelium (Pomin, V.H., Sharp, J.S., et al. 2010). Nevertheless, the number of local probes (one nitrogen per disaccharide unit), the partial labeling from glutamine incorporation and the exchangeable amine protons limits the use of this isotope for HS characterization. The usage of more specific HS precursors could solve this issue making it possible to investigate HS structure directly on the cell surface using for example solid state NMR.

Despite these limitations, the methods developed here should find ample applications. It is well known that different HS modification patterns, which occur in response to various environmental effects and during differentiation process, strongly influence development, physiology and pathology. At the level of cell culture, it is thus of wide importance to

characterize HS profile to predict how cells will respond to growth factor or cytokine stimuli. For example, for stem cells, where growth and differentiation are directly under the control of numerous HS binding cytokines, morphogens or growth factors (Kraushaar, D.C., Dalton, S., et al. 2013), or for cancer cells (Casu, B., Naggi, A., et al. 2010), which proliferation rate also depends on such factors, our approach should be of large interest. One major aspect of our method is the ability to directly detect the HS IdoA level, a characteristic hardly quantifiable by currently available techniques. IdoA residues provide flexibility to the HS chain, facilitating interactions with proteins (Casu, B. and Lindahl, U. 2001) and as such are important for the regulatory properties of the polysaccharide. Regarding the different cellular systems used in the present study, our data showed that following change in the expression level of one of the biosynthetic enzyme, NDST1, several structural modifications can be characterized, well beyond that catalyzed by the targeted enzyme (*N*-sulfation). These analyses thus give insights into the functional relationships that exist between the different biosynthetic enzymes. We also observed that an increase in the level of epimerization (as in differentiated Caco-2 cells) was accompanied by an increase in IdoA sulfation. However, a decrease in epimerization (as found in the NDST^{-/-} MEFs) did not alter the level of IdoA sulfation, which remained at 65%. Although this would have to be confirmed using a larger number of cells of different types, this observation could suggest that the 2-*O*-sulfotransferase was not the limiting factor *per se* but potentially its association with the C5-epimerase. Finally, an advantage of generating ¹³C-labeled HS is that chemical shift perturbation can be monitored during a binding reaction, so that the region of the polymer involved in an interaction can be directly identified and this will be investigated in future work.

MATERIALS AND METHODS

Cell culture and metabolic labeling

HeLa and Caco-2 cells were obtained from the European Collection of Cell Cultures. Mouse embryonic fibroblast (MEF) cell-lines were generated from primary fibroblasts isolated from E14.5 wild-type and NDST1^{-/-} mouse embryos as described previously (Ringvall, M., Ledin, J., et al. 2000). Cells were grown at 37 °C under a 5% CO₂ atmosphere high in Dulbecco's modified Eagle's medium containing ¹³C-glucose (4.5 mg ml⁻¹), supplemented with 10 % heat-inactivated fetal calf serum, 2 mM Glutamax and, for Caco-2 cells, 1 % non-essential amino acids. Caco-2 cells were differentiated as described (Sadir, R., Imberty, A., et al. 2004).

GAG extraction and purification

Cells were detached with EDTA (Versene, life technologies) and counted. Cell associated GAGs were extracted from 2x10⁸ cells by trypsin treatment, loaded onto a DEAE-Sephacel column as described (Lortat-Jacob, H., Turnbull, J.E., et al. 1995) and eluted with a 0.15-1.5 M linear NaCl gradient. Desalted samples were freeze-dried and contaminating nucleic acids were eliminated by DNase I treatment (Roche, 20 000 U.ml⁻¹ for 24 h at 37°C in 0.1 M Tris-HCl buffer, pH 8.0 containing 0.15 M NaCl and 10 mM MgCl₂). The reaction was stopped by adding 10 mM EDTA and the sample was repurified as described above.

Disaccharide analysis

GAGs, isolated from a T75 culture flask were digested with chondroitinase ABC and the digestion products were removed by DEAE ion exchange chromatography. The remaining HS were treated with heparinase I, II, and III and analyzed, in triplicate, using reversed-phase ion-pair (RPIP) HPLC as described (Ledin, J., Staatz, W., et al. 2004).

Flow cytometry

Cells (5×10^5 at ~80 % of confluence) were detached with versene (Hela, MEF) or accutase (Caco-2), resuspended in ice-cold PBS containing 1mM EDTA, 1% (w/v) bovine serum albumin and incubated for 60 min at 4°C with either 10 $\mu\text{g} \cdot \text{ml}^{-1}$ of anti-HS antibody (10E4, Seikagaku Corporation) or anti-CS (clone CS-56, Sigma, dilution 1/100). Cells were then washed three times in the same buffer, incubated with FITC-conjugated anti-mouse immunoglobulins (1/300; Interchim), fixed in 4% (w/v) paraformaldehyde buffer and analyzed by flow cytometry (Macs Quant, Miltenyi).

NMR analysis

Samples were prepared in D₂O (Sigma) by two cycles of freeze drying/resuspension into 99.9% D₂O, and transferred into a 3 mm diameter NMR tube. NMR experiments were recorded at 37°C either on a Varian 600 MHz Direct Drive or 700MHz Bruker advance III AD spectrometers equipped with a triple resonance (¹H, ¹³C, ¹⁵N) cold probe including shielded z-gradients. The ¹H-¹³C HSQC spectra were acquired with 128 scans per fid. The resolutions in the ¹H and ¹³C dimension were about 6 Hz and 100 Hz, respectively. ¹H and ¹³C dimensions were centered on 4.66 ppm and 79.9 ppm respectively and the experimental time was of 14 hours on a 600 MHz spectrometer. Transfer delays of the ¹H-¹³C experiments were matched to a J_{CH} value of 145 Hz. For the anomeric signals of IdoA and GlcN, possible J_{CH} mismatches were not corrected during the quantification process. Data were processed with NMRpipe (Delaglio, F., Grzesiek, S., et al. 1995) and analyzed with CcpNmr (Vranken, W.F., Boucher, W., et al. 2005). Proton chemical shifts were reported with respect to the H₂O signal relative to DSS at 37 °C. The ¹³C chemical shifts were referenced indirectly using the X/¹H frequency ratios of the zero-point: 0.251449530 (¹³C) (Wishart, D.S., Bigam, C.G., et al.

1995). The different saccharides were quantified by integrating the volume of unambiguous correlations as proposed by Guerrini *et al.* (Guerrini, M., Bisio, A., et al. 2001). Limits of integration of the peaks were manually adjusted using a home-made python program extracting spectra data.

FUNDING

This work was supported by the ANR (ANR-13-BSV3-0010-04), the Grenoble Partnership for Structural Biology (PSB) and used the platforms of the Grenoble Instruct Centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL), [grant numbers FRISBI (ANR-10-INSB-05-02), GRAL (ANR-10-LABX-49-01)].

ACKNOWLEDGMENTS

The authors thank C. Laguri for helpful discussions. MP, co supervised by RS and PG, was supported by a grant from the doctoral school of the University Grenoble-Alpes.

REFERENCES

- Bulow HE, Hobert O. 2006. The molecular diversity of glycosaminoglycans shapes animal development. *Annual review of cell and developmental biology*, 22:375-407.
- Casu B, Lindahl U. 2001. Structure and biological interactions of heparin and heparan sulfate. *Advances in carbohydrate chemistry and biochemistry*, 57:159-206.
- Casu B, Naggi A, Torri G. 2010. Heparin-derived heparan sulfate mimics to modulate heparan sulfate-protein interaction in inflammation and cancer. *Matrix biology : journal of the International Society for Matrix Biology*, 29:442-452.
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR*, 6:277-293.
- Desai UR, Wang HM, Linhardt RJ. 1993. Specificity studies on the heparin lyases from *Flavobacterium heparinum*. *Biochemistry*, 32:8140-8145.
- Galeotti F, Volpi N. 2013. Novel reverse-phase ion pair-high performance liquid chromatography separation of heparin, heparan sulfate and low molecular weight-heparins disaccharides and oligosaccharides. *Journal of chromatography. A*, 1284:141-147.
- Gallagher JT, Walker A. 1985. Molecular distinctions between heparan sulphate and heparin. Analysis of sulphation patterns indicates that heparan sulphate and heparin are separate families of N-sulphated polysaccharides. *The Biochemical journal*, 230:665-674.
- Guerrini M, Bisio A, Torri G. 2001. Combined quantitative (1)H and (13)C nuclear magnetic resonance spectroscopy for characterization of heparin preparations. *Seminars in thrombosis and hemostasis*, 27:473-482.

- Guerrini M, Naggi A, Guglieri S, Santarsiero R, Torri G. 2005. Complex glycosaminoglycans: profiling substitution patterns by two-dimensional nuclear magnetic resonance spectroscopy. *Analytical biochemistry*, 337:35-47.
- Hagner-McWhirter A, Lindahl U, Li J. 2000. Biosynthesis of heparin/heparan sulphate: mechanism of epimerization of glucuronyl C-5. *The Biochemical journal*, 347 Pt 1:69-75.
- Huckerby TN, Nieduszynski IA, Giannopoulos M, Weeks SD, Sadler IH, Lauder RM. 2005. Characterization of oligosaccharides from the chondroitin/dermatan sulfates. ¹H-NMR and ¹³C-NMR studies of reduced trisaccharides and hexasaccharides. *The FEBS journal*, 272:6276-6286.
- Kraushaar DC, Dalton S, Wang L. 2013. Heparan sulfate: a key regulator of embryonic stem cell fate. *Biological chemistry*, 394:741-751.
- Kreuger J, Kjellen L. 2012. Heparan sulfate biosynthesis: regulation and variability. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*, 60:898-907.
- Langeslay DJ, Beecher CN, Naggi A, Guerrini M, Torri G, Larive CK. 2013. Characterizing the microstructure of heparin and heparan sulfate using N-sulfoglucosamine ¹H and ¹⁵N NMR chemical shift analysis. *Analytical chemistry*, 85:1247-1255.
- Langeslay DJ, Beni S, Larive CK. 2012. A closer look at the nitrogen next door: ¹H-¹⁵N NMR methods for glycosaminoglycan structural characterization. *Journal of magnetic resonance (San Diego, Calif. : 1997)*, 216:169-174.
- Ledin J, Staatz W, Li JP, Gotte M, Selleck S, Kjellen L, Spillmann D. 2004. Heparan sulfate structure in mice with genetically modified heparan sulfate production. *The Journal of biological chemistry*, 279:42732-42741.
- Lindahl U, Li JP. 2009. Interactions between heparan sulfate and proteins-design and functional implications. *International review of cell and molecular biology*, 276:105-159.

- Lortat-Jacob H, Turnbull JE, Grimaud JA. 1995. Molecular organization of the interferon gamma-binding domain in heparan sulphate. *The Biochemical journal*, 310 (Pt 2):497-505.
- Ori A, Wilkinson MC, Fernig DG. 2008. The heparanome and regulation of cell function: structures, functions and challenges. *Frontiers in bioscience : a journal and virtual library*, 13:4309-4338.
- Pomin VH, Sharp JS, Li X, Wang L, Prestegard JH. 2010. Characterization of glycosaminoglycans by 15N NMR spectroscopy and in vivo isotopic labeling. *Analytical chemistry*, 82:4078-4088.
- Ringvall M, Ledin J, Holmborn K, van Kuppevelt T, Ellin F, Eriksson I, Olofsson AM, Kjellen L, Forsberg E. 2000. Defective heparan sulfate biosynthesis and neonatal lethality in mice lacking N-deacetylase/N-sulfotransferase-1. *The Journal of biological chemistry*, 275:25926-25930.
- Sadir R, Imberty A, Baleux F, Lortat-Jacob H. 2004. Heparan sulfate/heparin oligosaccharides protect stromal cell-derived factor-1 (SDF-1)/CXCL12 against proteolysis induced by CD26/dipeptidyl peptidase IV. *The Journal of biological chemistry*, 279:43854-43860.
- Sambuy Y, De Angelis I, Ranaldi G, Scarino ML, Stamatii A, Zucco F. 2005. The Caco-2 cell line as a model of the intestinal barrier: influence of cell and culture-related factors on Caco-2 cell functional characteristics. *Cell biology and toxicology*, 21:1-26.
- Sarrazin S, Lamanna WC, Esko JD. 2011. Heparan sulfate proteoglycans. *Cold Spring Harb Perspect Biol*, 3:1-33.
- Shi X, Zaia J. 2009. Organ-specific heparan sulfate structural phenotypes. *The Journal of biological chemistry*, 284:11806-11814.
- van Kuppevelt TH, Dennissen MA, van Venrooij WJ, Hoet RM, Veerkamp JH. 1998. Generation and application of type-specific anti-heparan sulfate antibodies using phage

1
2
3 display technology. Further evidence for heparan sulfate heterogeneity in the kidney. *The*
4
5 *Journal of biological chemistry*, 273:12960-12966.
6
7 Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL,
8
9 Ionides J, Laue ED. 2005. The CCPN data model for NMR spectroscopy: development of a
10
11 software pipeline. *Proteins*, 59:687-696.
12
13 Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD.
14
15 1995. ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *Journal of*
16
17 *biomolecular NMR*, 6:135-140.
18
19 Zaia J. 2013. Glycosaminoglycan glycomics using mass spectrometry. *Molecular & cellular*
20
21 *proteomics : MCP*, 12:885-892.
22
23 Zhang Z, McCallum SA, Xie J, Nieto L, Corzana F, Jimenez-Barbero J, Chen M, Liu J,
24
25 Linhardt RJ. 2008. Solution structures of chemoenzymatically synthesized heparin and its
26
27 precursors. *Journal of the American Chemical Society*, 130:12998-13007.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

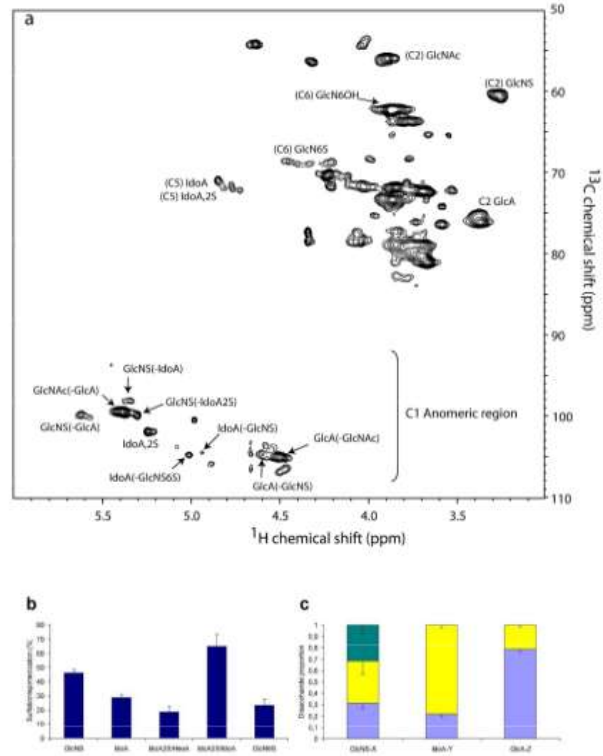
FIGURE LEGENDS

Figure 1: Hela cell HS characteristics. (a) 2D ^1H - ^{13}C HSQC spectrum of HS extracted from 1.5×10^8 cells grown in ^{13}C -glucose containing medium. Selected signals are labeled based on the assignments of Guerrini *et al.* (Guerrini, M., Naggi, A., et al. 2005) and used for cross-peaks integration. (b) Sulfation and epimerization (IdoA/GlcA) profile of Hela cell HS. (c) Disaccharide proportion of Hela HS in which GlcNS(\pm 6S) was associated to X = GlcA (blue), IdoA (Yellow) or IdoA2S (green); IdoA to Y = GlcNS (blue) or GlcNS6S (Yellow); GlcA linked to Z = GlcNAc (blue) or GlcNS (yellow).

Figure 2: Comparison of WT- and NDST1 KO- MEF HS characteristics. (a) Sulfation and epimerization profile of WT- (orange) and NDST1 KO- (green) MEFs. (b) Disaccharide proportion of MEF (WT and KO) HS in which GlcNS(\pm 6S) was associated to X = GlcA (blue), IdoA (Yellow) or IdoA2S (green); IdoA to Y = GlcNS (blue) or GlcNS6S (Yellow); GlcA linked to Z = GlcNAc (blue) or GlcNS (yellow).

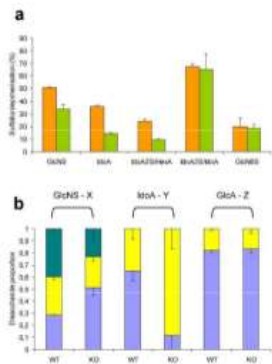
Figure 3: Comparison of non-differentiated and differentiated Caco-2 cell HS. (a) Sulfation and epimerization profile of non-differentiated- (orange) and differentiated- (green) Caco-2 cells. (b) Disaccharide proportion of non-differentiated- (Non-Diff) and differentiated- (Diff) Caco-2 cell HS in which GlcNS(\pm 6S) was associated to X = GlcA (blue), IdoA (Yellow) or IdoA2S (green); IdoA to Y = GlcNS (blue) or GlcNS6S (Yellow); GlcA linked to Z = GlcNAc (blue) or GlcNS (yellow).

Figure 1



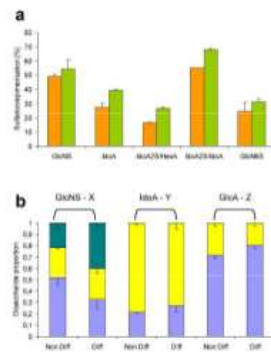
Hela cell HS characteristics
279x361mm (300 x 300 DPI)

Figure 2



Comparison of WT- and NDST1 KO- MEF HS characteristics
279x361mm (300 x 300 DPI)

Figure 3



Comparison of non-differentiated and differentiated Caco-2 cell HS
279x361mm (300 x 300 DPI)

SUPPORTING INFORMATION

Profiling sulfation/epimerization pattern of full-length heparan sulfate by NMR following cell culture ^{13}C -glucose metabolic labeling

Mathieu Pegéot^{1,2,3}, Rabia Sadir^{1,2,3}, Inger Eriksson⁴, Lena Kjellen⁴, Jean-Pierre Simorre^{1,2,3}, Pierre Gans^{1,2,3}, Hugues Iortat-Jacob^{1,2,3}

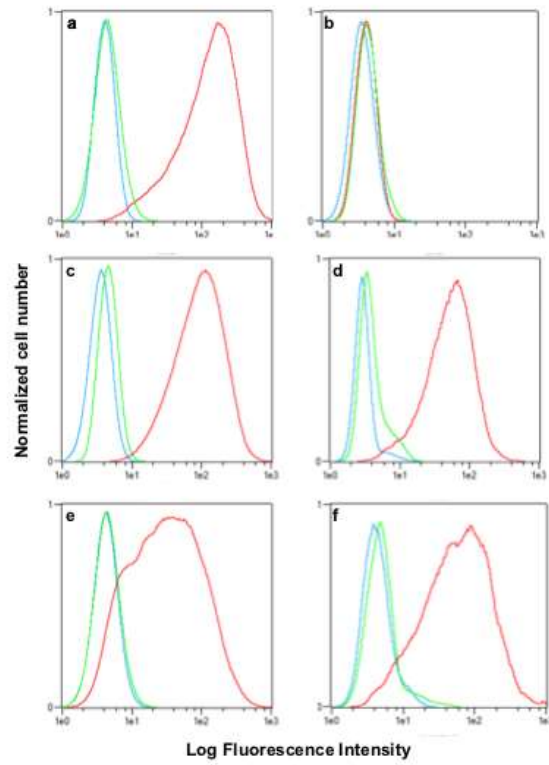
1 Univ. Grenoble Alpes, Institut de Biologie Structurale (IBS), F-38027 Grenoble, France

2 CNRS, IBS, F-38027 Grenoble, France

3 CEA, DSV, IBS, F-38027 Grenoble, France

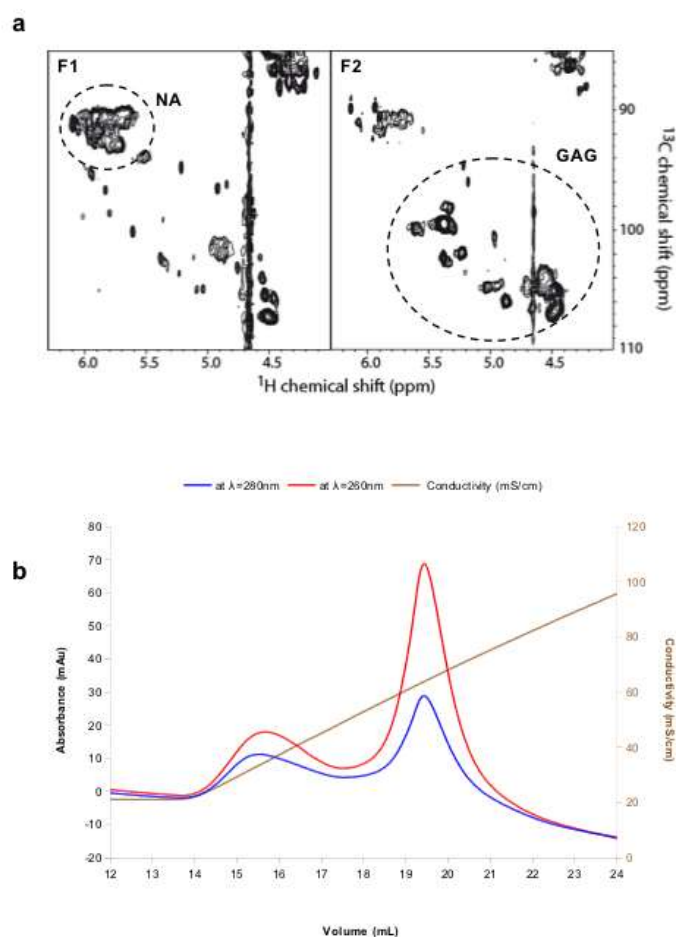
4 Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden.

Supplementary Figure 1



Flow cytometry analysis of cell associated HS and CS. (a) HeLa, (b) trypsin treated HeLa, (c) WT MEF, (d) KO MEF, (e) non-differentiated Caco-2 and (f) differentiated Caco-2 cells were incubated with 10 $\mu\text{g/ml}$ of anti-HS (red curves), anti-CS (green curves) or isotype control IgM (blue curves) antibody, then with FITC-conjugated anti-mouse immunoglobulins and analyzed by flow cytometry.

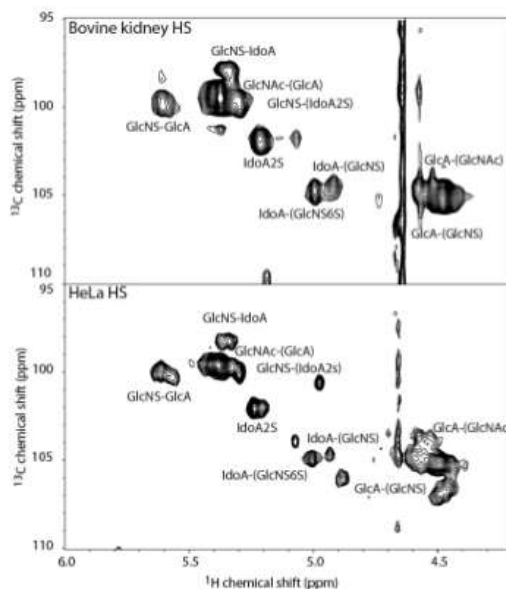
Supplementary Figure 2



GAGs preparation for NMR analysis. GAGs from 1.5×10^8 HeLa cells were extracted by extensive trypsin digestion and fractionated by DEAE-Sephacel chromatography as described (Lortat-Jacob, H., Turnbull, J.E., et al. 1995) (a) Eluted materials resolved into two overlapping fractions (F1 and F2) showing NMR signals predominantly attributed to nucleic acids^a (NA) or to HS/CS (GAG). (b) F2 material was further cleaned by DNase treatment and DEAE chromatography, freeze dried and resuspended into D₂O for NMR analysis.

^a Biological Magnetic Resonance Data Bank" (A Repository for Data from NMR Spectroscopy on Proteins, Peptides, Nucleic Acids, and other Biomolecules)

Supplementary Figure 3

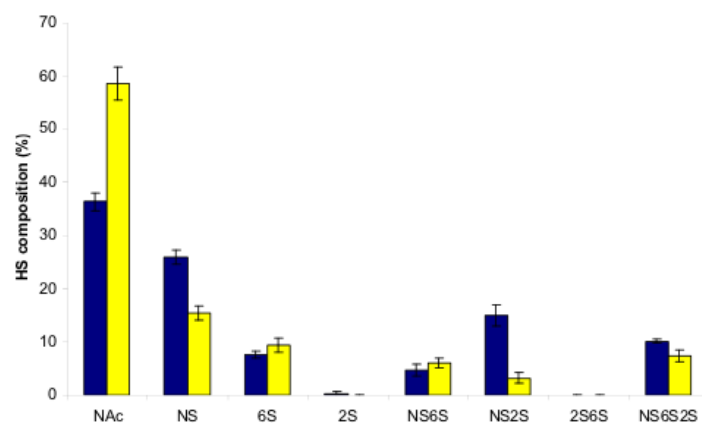


Comparison of the ^{13}C - ^1H HSQC spectra of HS from bovine kidney and Hela cells.

Bovine kidney HS (2 mg from Sigma-Aldrich, reference H7640) and Hela cell extracted HS were analyzed using identical NMR experiments recorded at 37°C on a Varian 600 MHz Direct Drive spectrometer equipped with a triple resonance (^1H , ^{13}C , ^{15}N) cold probe including shielded z-gradients. The ^{13}C - ^1H HSQC spectra were acquired in 200 t1 increments for each 128 scans for the HeLa HS sample and 120 t1 increments for each 180 scans for the bovine kidney HS sample (experimental times of 15.75 hours and 13.25 hours, respectively). Estimation of the Hela HS amount was made by comparison of the sum of the integrated volumes of the GlcNAc and GlcNS C2/H2 cross-peaks with the sum of the same cross-peaks of the 2 mg bovine kidney HS, taking in account the experimental time ratio. This comparison indicated that $\sim 10\ \mu\text{g}$ of ^{13}C -labeled HS were obtained from 2×10^8 Hela cells. This amount is consistent with the expected amount of HS synthesized by epithelial cells, assuming that each can produce 10^6 HS chains (Yanagishita, M. and Hascall, V.C. 1992) of 70-100 disaccharides (see supporting figure S5) having an average molecular mass of 500, thus a total mass of 12–16 μg .

Supplementary Figure 4

a

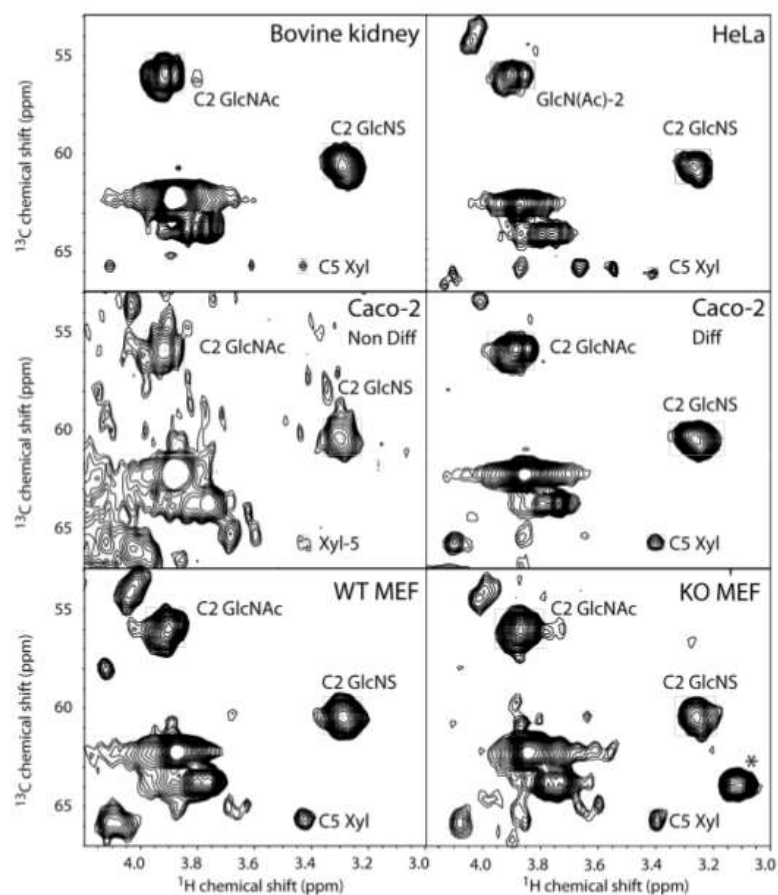


b

	WT		KO	
	Dp2 analysis	NMR	Dp2 analysis	NMR
N-sulfation	55.8 ± 2	50.7 ± 0.3	32 ± 2.4	33.7 ± 3.8
Epimerisation	na*	36 ± 1.1	na	14.37 ± 0.9
HexA 2-O-sulfation	25.4 ± 2.7	20.1 ± 1.5	10.6 ± 1.9	9.3 ± 1.2
IdoA 2-O-Sulfation	na	67.5 ± 2	na	65 ± 12
6-O-Sulfation	22.4 ± 2.7	24.1 ± 0.8	22.8 ± 3.2	18.4 ± 2.9
Sulfation/disaccharide	1.03 ± 0.03	0.95	0.65 ± 0.07	0.61

Disaccharide composition of MEF HS. (a) The composition of HS from WT- (blue) and KO- (yellow) MEFs was analyzed by RPIP HPLC(Ledin, J., Staatz, W., et al. 2004) and quantified against known amounts of standard disaccharides after heparinase mediated depolymerisation. (b) These data were used to calculate the level of N-, 6-*O* and 2-*O* sulfations and were compared to that obtained with the NMR based approach. *na: non analyzable with this technique

Supplementary Figure 5



Plots of the H2/C2 GlcNAc, GlcNS and H5/C5 xylose cross-peaks region of the ^{13}C - ^1H HSQC spectra of different HS samples. Upper left panel: Bovine kidney HS; Upper right panel: HeLa cells HS; Middle left panel: Undifferentiated Caco-2 cells HS; Middle right panel: differentiated Caco-2 cells HS; Lower left panel: wild-type MEF cells HS; Lower right panel: NDST1-KO cells HS. All spectra were plotted at 10% of the maximal intensity of the H2/C2 GlcN cross-peaks except for the boxed H5/C5 xylose parts for HeLa cells which were plotted at 5% of the maximal intensity. Estimations of the chain lengths (in disaccharide units) were made by comparison of the volume of the H5/C5 xylose cross-peaks with the sum of the volume of the H2/C2 cross-peaks. (Iacomini, M., Casu, B., et al. 1999)

REFERENCES

Iacomini M, Casu B, Guerrini M, Naggi A, Pirola A, Torri G. 1999. "Linkage region" sequences of heparins and heparan sulfates: detection and quantification by nuclear magnetic resonance spectroscopy. *Analytical biochemistry*, 274:50-58.

Ledin J, Staatz W, Li JP, Gotte M, Selleck S, Kjellen L, Spillmann D. 2004. Heparan sulfate structure in mice with genetically modified heparan sulfate production. *The Journal of biological chemistry*, 279:42732-42741.

Lortat-Jacob H, Turnbull JE, Grimaud JA. 1995. Molecular organization of the interferon gamma-binding domain in heparan sulphate. *The Biochemical journal*, 310 (Pt 2):497-505.

Yanagishita M, Hascall VC. 1992. Cell surface heparan sulfate proteoglycans. *The Journal of biological chemistry*, 267:9451-9454.

Bibliographie

- Abeijon C, Mandon EC, Hirschberg C. 1997. Transporters of nucleotide sugars, nucleotide sulfate and ATP in the Golgi apparatus. *TIBS*, pp. 203–7
- Abrink M, Grujic M, Pejler G. 2004. Serglycin is essential for maturation of mast cell secretory granule. *J. Biol. Chem.* 279(39):40897–905
- Aikawa J, Grobe K, Tsujimoto M, Esko JD. 2001. Multiple isozymes of heparan sulfate/heparin GlcNAc N-deacetylase/GlcN N-sulfotransferase. Structure and activity of the fourth member, NDST4. *J. Biol. Chem.* 276(8):5876–82
- Aikawa J -i., Esko JD. 1999. Molecular Cloning and Expression of a Third Member of the Heparan Sulfate/Heparin GlcNAcN-Deacetylase/ N-Sulfotransferase Family. *J. Biol. Chem.* 274(5):2690–95
- Al-Assaf S, Meadows J, Phillips GO, Williams PA, Parsons BJ. 2000. The effect of hydroxyl radicals on the rheological performance of hylan and hyaluronan. *Int. J. Biol. Macromol.* 27(5):337–48
- Almond A, Deangelis PL, Blundell CD. 2006. Hyaluronan: the local solution conformation determined by NMR and computer modeling is close to a contracted left-handed 4-fold helix. *J. Mol. Biol.* 358(5):1256–69
- Alvarez-Dominguez C, Vazquez-Boland J-A, Carrasco-Marin E, Lopez-Mato P, Leyva-Cobian F. 1997. Host cell heparan sulfate proteoglycans mediate attachment and entry of *Listeria monocytogenes*, and the listerial surface protein ActA is involved in heparan sulfate receptor recognition. *Infect. Immun.* 65(1):78–88
- Angulo J, Hricovíni M, Gairi M, Guerrini M, Paz JL De, et al. 2005. Dynamic properties of biologically active synthetic heparin-like hexasaccharides. *Glycobiology.* 15(10):1008–15
- Antonias KN, Fraser JR, Muirden KD. 1973. Distribution of biologically labelled radioactive hyaluronic acid injected into joints. *Ann. Rheum. Dis.* 32(2):103–11
- Aruffo A, Stamenkovic I, Melnick M, Underhill CB, Seed B. 1990. CD44 is the principal cell surface receptor for hyaluronate. *Cell.* 61(7):1303–13
- Arvan P, Castle D. 1998. Sorting and storage during secretory granule biogenesis: looking backward and looking forward. *Biochem J*, pp. 593–610
- Atkins EDT, Laurent TC. 1973. X-Ray-Diffraction Patterns from Chondroitin 4-Sulphate, Dermatan Sulphate and Heparan Sulphate. *Biochem J.* 133:605–6
- Atkins EDT, Nieduszynski I a, Horner AA. 1974. Crystallization of Macromolecular Heparin. *Biochem J.* 143:251–52
- Bai X, Zhou D, Brown JR, Crawford BE, Hennet T, Esko JD. 2001. Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth

- member of the beta 1,3-galactosyltransferase family (beta 3GalT6). *J. Biol. Chem.* 276(51):48189–95
- Bao X, Nishimura S, Mikami T, Yamada S, Itoh N, Sugahara K. 2004. Chondroitin sulfate/dermatan sulfate hybrid chains from embryonic pig brain, which contain a higher proportion of L-iduronic acid than those from adult pig brain, exhibit neuritogenic and growth factor binding activities. *J. Biol. Chem.* 279(11):9765–76
- Bar-Ner M, Eldor A, Wasserman L, Matzner Y, Cohen IR, et al. 1987. Inhibition of Heparanase-Mediated Degradation of Extracellular Matrix Heparan Sulfate by Non-anticoagulant Heparin Species. *Blood.* 70(2):551–57
- Bathe M, Rutledge GC, Grodzinsky AJ, Tidor B. 2005. A coarse-grained molecular model for glycosaminoglycans: application to chondroitin, chondroitin sulfate, and hyaluronic acid. *Biophys. J.* 88(6):3870–87
- Beauvais DM, Rapraeger AC. 2004. Syndecans in tumor cell adhesion and signaling. *Reprod. Biol. Endocrinol.* 2(3):1–12
- Bécharde D, Gentina T, Delehedde M, Scherpereel a, Lyon M, et al. 2001. Endocan is a novel chondroitin sulfate/dermatan sulfate proteoglycan that promotes hepatocyte growth factor/scatter factor mitogenic activity. *J. Biol. Chem.* 276(51):48341–49
- Bernfield M, Kokenyesi R, Kato M, Hinkes MT, Spring J, et al. 1992. Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. *Annu. Rev. Cell Biol.* 8:365–93
- Bezakova G, Ruegg M a. 2003. New insights into the roles of agrin. *Nat. Rev. Mol. Cell Biol.* 4(4):295–308
- Bie H, Yin J, He X, Kermode AR, Goddard-Borger ED, et al. 2013. Insights into mucopolysaccharidosis I from the structure and action of α -L-iduronidase. *Nat. Chem. Biol.* 9(11):739–45
- Bienkowski MJ, Conrad HE. 1985. Structural characterization of the oligosaccharides formed by depolymerization of heparin with nitrous acid. *J. Biol. Chem.* 260(1):356–65
- Bishop JR, Schuksz M, Esko JD. 2007. Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature.* 446(7139):1030–37
- Bitter T, Muir H. 1962. A Modified Uronic Acid Carbazole Reaction. *Anal. Biochem.* 4:330–34
- Blundell CD, DeAngelis PL, Day AJ, Almond A. 2004. Use of ^{15}N -NMR to resolve molecular details in isotopically-enriched carbohydrates: sequence-specific observations in hyaluronan oligomers up to decasaccharides. *Glycobiology.* 14(11):999–1009
- Bodenhausen G, Ruben DJ. 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* 69(1):185–89

- Borsig L, Vlodavsky I, Ishai-michaeli R, Torri G, Vismara E. 2011. Sulfated Hexasaccharides Attenuate Metastasis by Inhibition of P-selectin and Heparanase. *Neoplasia*. 13(5):445–52
- Bose B, Zhao S, Stenutz R, Cloran F, Bondo PB, et al. 1998. Three-Bond C - O - C - C Spin-Coupling Constants in Carbohydrates : Development of a Karplus Relationship. *J Am Chem Soc*. 120(43):11158–73
- Bradbury EJ, Moon LDF, Popat RJ, King VR, Bennett GS, et al. 2002. Chondroitinase ABC promotes functional recovery after spinal cord injury. *Nature*. 416(6881):636–40
- Brissett NC, Perkins SJ. 1998. Conserved basic residues in the C-type lectin and short complement repeat domains of the G3 region of proteoglycans. *Biochem J*. 329:415–24
- Brooks BR, Brooks III CL, MacKerell Jr AD, Nilsson L, Petrella RJ, et al. 2009. CHARMM: The Biomolecular Simulation Program. *J Comput Chem*. 30(10):1545–1614
- Bruckner P. 2010. Suprastructures of extracellular matrices: paradigms of functions controlled by aggregates rather than molecules. *Cell Tissue Res*. 339(1):7–18
- Bulow HE, Hobert O. 2004. Differential Sulfations and Epimerization Define Heparan Sulfate Specificity in Nervous System Development. *Neuron*. 41:723–36
- Busse M, Feta A, Presto J, Wilén M, Grønning M, et al. 2007. Contribution of EXT1, EXT2, and EXTL3 to heparan sulfate chain elongation. *J. Biol. Chem*. 282(45):32802–10
- Cael JJ, Winter WT, Arnott S. 1978. Calcium chondroitin 4-sulfate: molecular conformation and organization of polysaccharide chains in a proteoglycan. *J. Mol. Biol*. 125(1):21–42
- Calabro a., Midura R, Wang a., West L, Plaas a., Hascall VC. 2001. Fluorophore-assisted carbohydrate electrophoresis (FACE) of glycosaminoglycans. *Osteoarthr. Cartil*. 9:S16–S22
- Capila I, Linhardt RJ. 2002. Heparin - Protein Interactions. *Angew. Chemie*. 41:390–412
- Caplan MJ, Stow JL, Newman AP, Madri J, Clarke Anderson H, et al. 1987. Dependence on pH of polarized sorting of secreted proteins. *Nature*. 329:632–35
- Capurro MI, Xu P, Shi W, Li F, Jia A, Filmus J. 2008. Glypican-3 inhibits Hedgehog signaling during development by competing with patched for Hedgehog binding. *Dev. Cell*. 14(5):700–711
- Cardin AD, Weintraub HJ. 1989. Molecular Modeling of protein-glycosaminoglycan interactions. *Arterioscler. Thromb. Vasc. Biol*. 9(1):21–32
- Carlos TM, Schwartz BR, Kovach NL, Yee E, Rosa M, et al. 1990. Vascular cell adhesion molecule-1 mediates lymphocyte adherence to cytokine-activated cultured human endothelial cells. *Blood*. 76(5):965–70

- Carruthers VB, Håkansson S, Olivia K, Sibley LD, Giddings OK. 2000. Toxoplasma gondii Uses Sulfated Proteoglycans for Substrate and Host Cell Attachment. *Infect. Immun.* 68(7):4005–11
- Casu B, Choay J, Ferro DR, Gatti G, Jacquinet J-C, et al. 1986. Controversial glycosaminoglycan conformations. *Nature.* 322:215–16
- Casu B, Naggi A, Torri G. 2010. Heparin-derived heparan sulfate mimics to modulate heparan sulfate-protein interaction in inflammation and cancer. *Matrix Biol.* 29:442–52
- Casu B, Oreste P, Torri G, Zoppetti G, Choay J, et al. 1981. The structure of heparin oligosaccharide fragments with high anti-(factor Xa) activity containing the minimal antithrombin III-binding sequence. *Biochem J.* 197:599–609
- Chakravarti S, Magnuson T, Lass JH, Jepsen KJ, LaMantia C, Carroll H. 1998. Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican. *J. Cell Biol.* 141(5):1277–86
- Chakravarti S, Petroll WM, Hassell JR, Jester J V, Lass JH, et al. 2000. Corneal opacity in lumican-null mice: defects in collagen fibril structure and packing in the posterior stroma. *Invest. Ophthalmol. Vis. Sci.* 41(11):3365–73
- Chen P, Abacherli LE, Nadler ST, Wang Y, Li Q, Parks WC. 2009. MMP7 shedding of syndecan-1 facilitates re-epithelialization by affecting alpha(2)beta(1) integrin activation. *PLoS One.* 4(8):e6565
- Chen X-D, Fisher LW, Robey PG, Young MF. 2004. The small leucine-rich proteoglycan biglycan modulates BMP-4-induced osteoblast differentiation. *FASEB J.* 18(9):948–58
- Chen Y, Götte M, Liu J, Park PW. 2008. Microbial subversion of heparan sulfate proteoglycans. *Mol. Cells.* 26(5):415–26
- Chen Y, Hayashida A, Bennett AE, Hollingshead SK, Park PW. 2007. Streptococcus pneumoniae sheds syndecan-1 ectodomains through ZmpC, a metalloproteinase virulence factor. *J. Biol. Chem.* 282(1):159–67
- Chihara E, Tsukada-fujioka T, Tsukahara I. 1984. Synthesis of glycosaminoglycans by chick retinal pigment epithelium in vitro. *Graefe's Arch Clin Exp Ophthalmol.* 222:45–50
- Choi S, Lee E, Kwon S, Park H, Yi JY, et al. 2005. Transmembrane domain-induced oligomerization is crucial for the functions of syndecan-2 and syndecan-4. *J. Biol. Chem.* 280(52):42573–79
- Christner JE, Distler JJ, Jourdian GW. 1979. Biosynthesis of keratan sulfate: purification and properties of a galactosyltransferase from bovine cornea. *Arch. Biochem. Biophys.* 192(2):548–58
- Cohlberg JA, Li J, Uversky VN, Fink AL. 2002. Heparin and Other Glycosaminoglycans Stimulate the Formation of Amyloid Fibrils from alpha-Synuclein in Vitro. *Biochemistry.* 41:1502–11

- Colebrooke SA, Blundell CD, Deangelis PL, Campbell ID, Almond A. 2005. Exploiting the carboxylate chemical shift to resolve degenerate resonances in spectra of ^{13}C -labelled glycosaminoglycans. *Magn. Reson. Chem.* 43:805–15
- Connell BJ, Baleux F, Coic Y, Clayette P, Bonnaffe D, Lortat-Jacob H. 2012. Article A Synthetic Heparan Sulfate-Mimetic Peptide Conjugated to a Mini CD4 Displays Very High Anti- HIV-1 Activity Independently of Coreceptor Usage. *Chem. Biol.* 19:131–39
- Cortijo-Arellano M, Ponce J, Durany N, Cladera J. 2008. Amyloidogenic properties of the prion protein fragment PrP (185–208): Comparison with Alzheimer's peptide Abeta(1–28), influence of heparin and cell toxicity. *Biochem. Biophys. Res. Commun.* 368:238–42
- Couchman JR. 2010. Transmembrane signaling proteoglycans. *Annu. Rev. Cell Dev. Biol.* 26:89–114
- Cowman MK, Feder-davis J, Hittner DM. 2001. ^{13}C NMR Studies of Hyaluronan . 2 . Dependence of Conformational Dynamics on Chain Length and Solvent. *Macromolecules.* 34:110–15
- Cros S, Petitou M, Sizun P, Pérez S, Imberty A. 1997. Combined NMR and Molecular Modeling Study of an Iduronic Acid-Containing Trisaccharide Related to Antithrombotic Heparin Fragments. *Bioorg. Med. Chem.* 5(7):1301–9
- Daniele A, Tomanin R, Villani GRD, Zacchello F, Scarpa M, Natale P Di. 2002. Uptake of recombinant iduronate-2-sulfatase into neuronal and glial cells in vitro. *Biochim. Biophys. Acta.* 1588:203–9
- Davidson E, Meyer K. 1955. Structural studies on chondroitin sulfuric acid. II. The glucuronic linkage. *J Am Chem Soc.* 77:4796–98
- Day AJ, Sheehan JK. 2001. Hyaluronan: polysaccharide chaos to protein organisation. *Curr. Opin. Struct. Biol.* 11(5):617–22
- DeAngelis PL, Jing W, Graves M V., Burbank DE, Van Etten JL. 1997. Hyaluronan Synthase of Chlorella Virus PBCV-1. *Science (80-.).* 278(5344):1800–1803
- DeAngelis PL, Weigel PH. 1994. Immunochemical confirmation of the primary structure of streptococcal hyaluronan synthase and synthesis of high molecular weight product by the recombinant enzyme. *Biochemistry.* 33(31):9033–39
- Desai UR, Wang H-M, Ampofo SA, Linhardt RJ. 1993. Oligosaccharide Composition of Heparin and Low-Molecular-Weight Heparins by Capillary Electrophoresis. *Anal. Biochem.* 213:120–27
- Dhoot GK, Gustafsson MK, Ai X, Sun W, Standiford DM, Emerson CP. 2001. Regulation of Wnt signaling and embryo patterning by an extracellular sulfatase. *Science (80-.).* 293(5535):1663–66
- Edwards IJ. 2012. Proteoglycans in prostate cancer. *Nat. Rev. Urol.* 9(4):196–206

- Elfenbein A, Simons M. 2013. Syndecan-4 signaling at a glance. *J. Cell Sci.* 126(Pt 17):3799–3804
- Esko JD, Selleck SB. 2002. Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem.* 71:435–71
- Esko JD, Zhang L. 1996. Influence of core protein sequence on glycosaminoglycan assembly. *Curr. Opin. Struct. Biol.* 5:663–70
- Faham S, Hileman RE, Fromm JR, Linhardt RJ, Rees DC. 1996. Heparin structure and interactions with basic fibroblast growth factor. *Science.* 271(5252):1116–20
- Fan G, Xiao L, Cheng L, Wang X, Sun B, Hu G. 2000. Targeted disruption of NDST-1 gene leads to pulmonary hypoplasia and neonatal respiratory distress in mice. *FEBS Lett.* 467(1):7–11
- Farach-Carson MC, Warren CR, Harrington D a, Carson DD. 2014. Border patrol: insights into the unique role of perlecan/heparan sulfate proteoglycan 2 at cell and tissue borders. *Matrix Biol.* 34:64–79
- Ferro DR, Provasoli A, Ragazzi M, Casu B, Torri G, et al. 1990. Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences. *Carbohydr. Res.* 195:157–67
- Feyerabend TB, Li J, Lindahl U, Rodewald H. 2006. Heparan sulfate C5-epimerase is essential for heparin biosynthesis in mast cells. *Nat. Chem. Biol.* 2(4):195–96
- Filmus J. 2001. Glypicans in growth control and cancer. *Glycobiology.* 11(3):19R–23R
- Filmus J, Capurro M, Rast J. 2008. Glypicans. *Genome Biol.* 9(5):224
- Fogh J, Fogh JM, Orfeo T. 1977. One hundred and twenty-seven cultured human tumor cell lines producing tumors in nude mice. *J Natl Cancer Inst.* 59(1):221–26
- Food and Drug Administration.
<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291390.pdf>
- Fransson L a, Belting M, Jönsson M, Mani K, Moses J, Oldberg a. 2000. Biosynthesis of decorin and glypican. *Matrix Biol.* 19(4):367–76
- Fransson L-A, Silverberg I, Carlstedt I. 1985. Structure of the Heparan Sulfate-Protein Linkage Region. *J. Biochem.* 260(27):14722–26
- Fried M, Domingo GJ, Gowda CD, Mutabingwa TK, Duffy PE. 2006. Plasmodium falciparum: chondroitin sulfate A is the major receptor for adhesion of parasitized erythrocytes in the placenta. *Exp. Parasitol.* 113(1):36–42
- Fuchs K, Hippe A, Schmaus A, Homey B, Sleeman JP. 2013. Opposing effects of high- and low-molecular weight hyaluronan on CXCL12-induced CXCR4 signaling depend on CD44. *Cell Death Dis.* 4(10):e819–11

- Fuki I V, Meyer ME, Williams KJ. 2000. Transmembrane and cytoplasmic domains of syndecan mediate a multi-step endocytic pathway involving detergent-insoluble membrane rafts. *Biochem J.* 351:607–12
- Fukuta M, Inazawa J, Torii T, Tsuzuki K, Shimada E, Habuchi O. 1997. Molecular Cloning and Characterization of Human Keratan Sulfate Gal-6-Sulfotransferase. *J. Biol. Chem.* 272(51):32321–28
- Fukuzumi M, Maruyama S, Sano M, Fukui S. 2001. Comparison of the expression of cell surface poly-N-acetyllactosamine-type oligosaccharides in PC12 cells with those in its variant PC12D. *Glycobiology.* 11(6):481–94
- Funderburgh JL. 2000. Keratan sulfate: structure, biosynthesis, and function. *Glycobiology.* 10(10):951–58
- Funderburgh JL. 2002. Keratan Sulfate Biosynthesis. *IUBMB Life.* 54(4):187–94
- Fuster MM, Esko JD. 2005. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat. Rev. Cancer.* 5(7):526–42
- Gallagher JT, Walker A. 1985. Molecular distinctions between heparan sulphate and heparin. *Biochem J.* 230:665–74
- Galli SJ, Nakae S, Tsai M. 2005. Mast cells in the development of adaptive immune responses. *Nat. Immunol.* 6(2):135–42
- Gehling JG, Rigby JK. 1996. Long Expected Sponges from the Neoproterozoic Ediacara Fauna of South Australia. *J. Paleontol.* 70(2):185–95
- George EL, Georges-Labouesse EN, Patel-King RS, Rayburn H, Hynes RO. 1993. Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. *Development.* 119(4):1079–91
- Gerweck LE, Seetharaman K. 1996. Cellular pH Gradient in Tumor versus Normal Tissue : Potential Exploitation for the Treatment of Cancer. *Cancer Res.* 56:1194–98
- Gingis-Velitski S, Zetser A, Flugelman MY, Vlodavsky I, Ilan N. 2004. Heparanase induces endothelial cell migration via protein kinase B/Akt activation. *J. Biol. Chem.* 279(22):23536–41
- Goedert M, Jakes R, Spillantini MG, Hasegawa M, Smith MJ, Crowther RA. 1996. Assembly of microtubule-associated protein tau into Alzheimer-like filaments induced by sulphated glycosaminoglycans. *Nature.* 383:550–53
- Greene DK, Tumova S, Couchman JR, Woods A. 2003. Syndecan-4 associates with alpha-actinin. *J. Biol. Chem.* 278(9):7617–23
- Grobe K, Ledin J, Ringvall M, Holmborn K, Forsberg E, et al. 2002. Heparan sulfate and development: differential roles of the N-acetylglucosamine N-deacetylase/N-sulfotransferase isozymes. *Biochim. Biophys. Acta.* 1573(3):209–15

- Guerrini M, Bisio a, Torri G. 2001. Combined quantitative (1)H and (13)C nuclear magnetic resonance spectroscopy for characterization of heparin preparations. *Semin. Thromb. Hemost.* 27(5):473–82
- Guerrini M, Naggi A, Guglieri S, Santarsiero R, Torri G. 2005. Complex glycosaminoglycans: profiling substitution patterns by two-dimensional nuclear magnetic resonance spectroscopy. *Anal. Biochem.* 337(1):35–47
- Gulberti S, Lattard V, Fondeur M, Jacquinet J-C, Mulliert G, et al. 2005. Phosphorylation and sulfation of oligosaccharide substrates critically influence the activity of human beta1,4-galactosyltransferase 7 (GalT-I) and beta1,3-glucuronosyltransferase I (GlcAT-I) involved in the biosynthesis of the glycosaminoglycan-protein I. *J. Biol. Chem.* 280(2):1417–25
- Guo BP, Brown EL, Dorward DW, Rosenberg LC, Höök M. 1998. Decorin-binding adhesins from *Borrelia burgdorferi*. *Mol. Microbiol.* 30(4):711–23
- Haasnoot CAG, de Leeuw FAAM, Altona C. 1980. THE RELATIONSHIP BETWEEN PROTON-PROTON COUPLING CONSTANTS AND SUBSTITUENT ELECTRONEGATIVITIES. AN EMPIRICAL GENERALIZATION OF THE KARPLUS EQUATION. *Tetrahedron.* 36:2783–92
- Habuchi H, Miyake G, Nogami K, Kuroiwa A, Matsuda Y, et al. 2003. Biosynthesis of heparan sulphate with diverse structures and functions: two alternatively spliced forms of human heparan sulphate 6-O-sulphotransferase-2 having different expression patterns and properties. *Biochem. J.* 371(Pt 1):131–42
- Habuchi H, Nagai N, Sugaya N, Atsumi F, Stevens RL, Kimata K. 2007. Mice deficient in heparan sulfate 6-O-sulfotransferase-1 exhibit defective heparan sulfate biosynthesis, abnormal placentation, and late embryonic lethality. *J. Biol. Chem.* 282(21):15578–88
- Habuchi H, Tanaka M, Habuchi O, Yoshida K, Suzuki H, et al. 2000. The Occurrence of Three Isoforms of Heparan Sulfate 6-O-Sulfotransferase Having Different Specificities for Hexuronic Acid Adjacent to the Targeted N-Sulfoglucosamine. *J. Biol. Chem.* 275(4):2859–68
- Harris NC, Davydova N, Roufail S, Paquet-Fifield S, Paavonen K, et al. 2013. The Propeptides of VEGF-D Determine Heparin Binding, Receptor Heterodimerization, and Effects on Tumor Biology. *J. Biol. Chem.* 288(12):8176–86
- Hempel N, How T, Dong M, Murphy SK, Fields T a, Blobel GC. 2007. Loss of betaglycan expression in ovarian cancer: role in motility and invasion. *Cancer Res.* 67(11):5231–38
- Higuchi T, Shimizu H, Fukuda T, Kawagoe S, Matsumoto J, et al. 2012. Enzyme replacement therapy (ERT) procedure for mucopolysaccharidosis type II (MPS II) by intraventricular administration (IVA) in murine MPS II. *Mol. Genet. Metab.* 107:122–28
- Hileman RE, Fromm JR, Weiler JM, Linhardt RJ. 1998. Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *BioEssays.* 20(2):156–67

- Hoffman P, Linker A, Meyer K. 1958. The acid mucopolysaccharides of connective tissue. III. The sulfate linkage. *Biochim Biophys Acta*. 30:184–85
- Holst CR, Bou-Reslan H, Gore BB, Wong K, Grant D, et al. 2007. Secreted sulfatases Sulf1 and Sulf2 have overlapping yet essential roles in mouse neonatal survival. *PLoS One*. 2(6):e575
- Homans SW. 1990. A molecular mechanical force field for the conformational analysis of oligosaccharides: comparison of theoretical and crystal structures of Man alpha 1-3Man beta 1-4GlcNAc. *Biochemistry*. 29(39):9110–18
- Hricovini M. 2011. Effect of Solvent and Counterions upon Structure and NMR Spin - Spin Coupling Constants in Heparin Disaccharide. *J. Phys. Chem. B*. 115:1503–11
- Hricovini M, Guerrini M, Torri G, Casu B. 1997. Motional properties of E. Coli polysaccharide K5 in aqueous solution analyzed by NMR relaxation measurements. *Carbohydr. Res*. 300:69–76
- Hricovlni M, Torri G. 1995. Dynamics in aqueous solutions of the pentasaccharide corresponding to the binding site of heparin for antithrombin III studied by NMR relaxation measurements. *Carbohydr. Res*. 268:159–75
- Huckerby TN, Nieduszynski IA. 1982. Proton chemical-shift assignments in the n.m.r spectra of heparan and heparin. *Carbohydr. Res*. 103:141–45
- Huckerby TN, Nieduszynski I a, Giannopoulos M, Weeks SD, Sadler IH, Lauder RM. 2005. Characterization of oligosaccharides from the chondroitin/dermatan sulfates. ¹H-NMR and ¹³C-NMR studies of reduced trisaccharides and hexasaccharides. *FEBS J*. 272(24):6276–86
- Humphries DE, Wong GW, Friend DS, Gurish MF, Qiu W, et al. 1999. Heparin is essential for the storage of specific granule proteases in mast cells. *Nature*. 400:769–72
- Iacomini M, Casu B, Guerrini M, Naggi a, Pirola a, Torri G. 1999. “Linkage region” sequences of heparins and heparan sulfates: detection and quantification by nuclear magnetic resonance spectroscopy. *Anal. Biochem*. 274(1):50–58
- Imberty a, Pérez S. 2000. Structure, conformation, and dynamics of bioactive oligosaccharides: theoretical approaches and experimental validations. *Chem. Rev*. 100(12):4567–88
- Imberty A, Lortat-Jacob H, Pérez S. 2007. Structural view of glycosaminoglycan-protein interactions. *Carbohydr. Res*. 342(3-4):430–39
- Iozzo V, Murdoch AD. 1996. Proteoglycans of the extracellular environment : clues from the gene and protein side offer novel perspectives in molecular diversity and function. *FASEB J*. 10(April):598–614
- Iozzo R V. 1998. Matrix proteoglycans: from molecular design to cellular function. *Annu. Rev. Biochem*. 67:609–52

- Iozzo R V. 2005. Basement membrane proteoglycans: from cellar to ceiling. *Nat. Rev. Mol. Cell Biol.* 6(8):646–56
- Iozzo R V, Sanderson RD. 2011. Proteoglycans in cancer biology, tumour microenvironment and angiogenesis. *J. Cell. Mol. Med.* 15(5):1013–31
- Iozzo R V., Moscatello DK, McQuillan DJ, Eichstetter I. 1999. Decorin Is a Biological Ligand for the Epidermal Growth Factor Receptor. *J. Biol. Chem.* 274(8):4489–92
- Itano N, Sawai T, Yoshida M, Lenas P, Yamada Y, et al. 1999. Three Isoforms of Mammalian Hyaluronan Synthases Have Distinct Enzymatic Properties. *J. Biol. Chem.* 274(35):25085–92
- IUPAC-IUB Joint Commission on Biochemical Nomenclature. 1983. Symbols for Specifying the Conformation of Polysaccharide Chains. *Eur J Biochem.* 131:5–7
- Izumikawa T, Koike T, Shiozawa S, Sugahara K, Tamura J, Kitagawa H. 2008. Identification of chondroitin sulfate glucuronyltransferase as chondroitin synthase-3 involved in chondroitin polymerization: chondroitin polymerization is achieved by multiple enzyme complexes consisting of chondroitin synthase family members. *J. Biol. Chem.* 283(17):11396–406
- Izumikawa T, Uyama T, Okuura Y, Sugahara K, Kitagawa H. 2007. Involvement of chondroitin sulfate synthase-3 (chondroitin synthase-2) in chondroitin polymerization through its interaction with chondroitin synthase-1 or chondroitin-polymerizing factor. *Biochem. J.* 403(3):545–52
- Jacobsson I, Lindahl U, Jensen JW, Roden L, Prihar H, Feingold DS. 1984. Biosynthesis of Heparin. Substrate Specificity of Heparosan N-Sulfate D-glucuronosyl 5-epimerase. *J. Biol. Chem.* 259(2):1056–63
- Jasnin M, van Eijck L, Marek Koza M, Peters J, Laguri C, et al. 2010. Dynamics of heparan sulfate explored by neutron scattering. *Phys. Chem. Chem. Phys.* 12:3360–62
- Jia J, Maccarana M, Zhang X, Beshpalov M, Lindahl U, Li J-P. 2009. Lack of L-iduronic acid in heparan sulfate affects interaction with growth factors and cell signaling. *J. Biol. Chem.* 284(23):15942–50
- Jin L, Abrahams JP, Skinner R, Petitou M, Pike RN, Carrell RW. 1997. The anticoagulant activation of antithrombin by heparin. *Proc. Natl. Acad. Sci. U. S. A.* 94(26):14683–88
- Kadler KE, Hill A, Canty-Laird EG. 2008. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. *Curr. Opin. Cell Biol.* 20(5):495–501
- Kalamajski S, Oldberg A. 2009. Homologous sequence in lumican and fibromodulin leucine-rich repeat 5-7 competes for collagen binding. *J. Biol. Chem.* 284(1):534–39
- Kalamajski S, Oldberg A. 2010. The role of small leucine-rich proteoglycans in collagen fibrillogenesis. *Matrix Biol.* 29(4):248–53

- Kaneiwa T, Mizumoto S, Sugahara K, Yamada S. 2010. Identification of human hyaluronidase-4 as a novel chondroitin sulfate hydrolase that preferentially cleaves the galactosaminidic linkage in the trisulfated tetrasaccharide sequence. *Glycobiology*. 20(3):300–309
- Karamanos NK, Vanky P, Tzanakakis GN, Hjerpe a. 1996. High performance capillary electrophoresis method to characterize heparin and heparan sulfate disaccharides. *Electrophoresis*. 17(2):391–95
- Karvinen S, Pasonen-Seppänen S, Hyttinen JMT, Pienimäki J-P, Törrönen K, et al. 2003. Keratinocyte growth factor stimulates migration and hyaluronan synthesis in the epidermis by activation of keratinocyte hyaluronan synthases 2 and 3. *J. Biol. Chem*. 278(49):49495–504
- Kendall FE, Heidelberger M, Dawson MH. 1937. A serologically inactive polysaccharide elaborated by mucoid strains of Group A Streptococcus. *J. Biochem.*, pp. 61–69
- Khan S, Fung KW, Rodriguez E, Patel R, Gor J, et al. 2013. The solution structure of heparan sulfate differs from that of heparin: implications for function. *J. Biol. Chem*. 288(39):27737–51
- Khan S, Gor J, Mulloy B, Perkins SJ. 2010. Semi-rigid solution structures of heparin by constrained X-ray scattering modelling: new insight into heparin-protein complexes. *J. Mol. Biol*. 395(3):504–21
- Khurana A, Belefard D, He X, Chien J, Shridhar V. 2013. Role of heparan sulfatases in ovarian and breast cancer. *Am. J. Cancer Res*. 3(1):34–45
- Kim B-T, Kitagawa H, Tanaka J, Tamura J, Sugahara K. 2003. In vitro heparan sulfate polymerization: crucial roles of core protein moieties of primer substrates in addition to the EXT1-EXT2 interaction. *J. Biol. Chem*. 278(43):41618–23
- Kirkpatrick C a, Knox SM, Staatz WD, Fox B, Lercher DM, Selleck SB. 2006. The function of a Drosophila glypican does not depend entirely on heparan sulfate modification. *Dev. Biol*. 300(2):570–82
- Kitagawa H. 1998. Molecular Cloning and Expression of Glucuronyltransferase I Involved in the Biosynthesis of the Glycosaminoglycan-Protein Linkage Region of Proteoglycans. *J. Biol. Chem*. 273(12):6615–18
- Kitagawa H, Izumikawa T, Uyama T, Sugahara K. 2003. Molecular cloning of a chondroitin polymerizing factor that cooperates with chondroitin synthase for chondroitin polymerization. *J. Biol. Chem*. 278(26):23666–71
- Klüttel M, Wight TN, Chan C, Hinek A, Wrana JL. 2005. Maintenance of chondroitin sulfation balance by chondroitin-4-sulfotransferase 1 is required for chondrocyte development and growth factor signaling during cartilage morphogenesis. *Development*. 132(17):3989–4003
- Knelson EH, Gaviglio AL, Nee JC, Starr MD, Nixon AB, et al. 2014a. Stromal heparan sulfate differentiates neuroblasts to suppress neuroblastoma growth. *J. Clin. Invest*. 124(7):3016–31

- Knelson EH, Nee JC, Blobel GC. 2014b. Heparan sulfate signaling in cancer. *Trends Biochem. Sci.* 39(6):277–88
- Knox SM, Whitelock JM. 2006. Perlecan: how does one molecule do so many things? *Cell. Mol. Life Sci.* 63(21):2435–45
- Kobayashi H, Suzuki M, Kanayama N, Nishida T, Takigawa M, Terao T. 2002. CD44 STIMULATION BY FRAGMENTED HYALURONIC ACID INDUCES UPREGULATION OF UROKINASE-TYPE PLASMINOGEN ACTIVATOR AND ITS RECEPTOR AND SUBSEQUENTLY FACILITATES INVASION OF HUMAN CHONDROSARCOMA CELLS. *Int J Cancer.* 102:379–89
- Kobayashi M, Sugumaran G, Liu J, Shworak NW, Silbert JE, Rosenberg RD. 1999. Molecular Cloning and Characterization of a Human Uranyl 2-Sulfotransferase That Sulfates Iduronyl and Glucuronyl Residues in Dermatan/Chondroitin Sulfate. *J. Biol. Chem.* 274(15):10474–80
- Kobayashi S, Morimoto K, Shimizu T, Takahashi M, Kurosawa H, Shirasawa T. 2000. Association of EXT1 and EXT2, Hereditary Multiple Exostoses Gene Products, in Golgi Apparatus. *Biochem. Biophys. Res. Commun.* 268(3):860–67
- Kobe B, Deisenhofer J. 1993. Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature.* 366(6457):751–56
- Kohda D, Morton CJ, Parkar a a, Hatanaka H, Inagaki FM, et al. 1996. Solution structure of the link module: a hyaluronan-binding domain involved in extracellular matrix stability and cell migration. *Cell.* 86(5):767–75
- Kolb M, Margetts PJ, Sime PJ, Gauldie J. 2001. Proteoglycans decorin and biglycan differentially modulate TGF-beta-mediated fibrotic responses in the lung. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 280(6):L1327–34
- Kolset S, Mann DM, Uhlin-Hansen L, Winberg J-O, Ruoslahti E. 1996. Serglycin-binding proteins in activated macrophages and platelets. *J. Leukoc. Biol.* 59(April):545–54
- Kolset SO, Prydz K, Pejler G. 2004. Intracellular proteoglycans. *Biochem. J.* 379(Pt 2):217–27
- Kolset SO, Tveit H. 2008. Serglycin--structure and biology. *Cell. Mol. Life Sci.* 65(7-8):1073–85
- Kraushaar DC, Dalton S, Wang L. 2013. Heparan Sulfate : a Key Regulator of Embryonic Stem Cell Fate. *Biol Chem.* 394(6):741–51
- Kreuger J, Kjellén L. 2012. Heparan Sulfate Biosynthesis : Regulation and Variability. *J. Histochem. Cytochem.* 60(12):898–907
- Kreuger J, Spillmann D, Li J, Lindahl U. 2006. Interactions between heparan sulfate and proteins: the concept of specificity. *J. Cell Biol.* 174(3):323–27
- Krukenberg CFW. 1884. Die chemischen Bestandteile des Knorpels. *Z. Biol.* 20:307

- Kulseth MA, Kolset SO, Ranheim T. 1999. Stimulation of serglycin and CD44 mRNA expression in endothelial cells exposed to TNF- α and IL-1 α . *Biochim. Biophys. Acta - Gen. Subj.* 1428(2-3):225–32
- Kusche-Gullberg M. 1998. Identification and Expression in Mouse of Two Heparan Sulfate Glucosaminyl N-Acetylase/N-Sulfotransferase Genes. *J. Biol. Chem.* 273(19):11902–7
- Kusche-Gullberg M, Kjellén L. 2003. Sulfotransferases in glycosaminoglycan biosynthesis. *Curr. Opin. Struct. Biol.* 13:605–11
- Laguri C, Sapay N, Simorre J, Brutscher B, Imberty A, et al. 2011. ¹³C-Labeled Heparan Sulfate Analogue as a Tool To Study Protein/Heparan Sulfate Interactions by NMR Spectroscopy: Application to the CXCL12a Chemokine. *J. Am. Chem. Soc.*
- Lai J-P, Sandhu DS, Yu C, Han T, Moser CD, et al. 2008. Sulfatase 2 up-regulates glypican 3, promotes fibroblast growth factor signaling, and decreases survival in hepatocellular carcinoma. *Hepatology.* 47(4):1211–22
- Lamant M, Smih F, Harmancey R, Philip-Couderc P, Pathak A, et al. 2006. ApoO, a novel apolipoprotein, is an original glycoprotein up-regulated by diabetes in human heart. *J. Biol. Chem.* 281(47):36289–302
- Lamb DJ, Wang HM, Mallis LM, Linhardt RJ. 1992. Negative ion fast-atom bombardment tandem mass spectrometry to determine sulfate and linkage position in glycosaminoglycan-derived disaccharides. *J Am Soc Mass Spectrom.* 3(8):797–803
- Laurent UBG, Laurent TC. 1981. On the origin of hyaluronate in blood. *Biochem Int.* 2:195–99
- Lesley J, Hascall VC, Tammi M, Hyman R. 2000. Hyaluronan binding by cell surface CD44. *J. Biol. Chem.* 275(35):26967–75
- Levene P, LaForge F. 1914. On chondroitin sulfuric acid. *J. Biol. Chem.* 18:123–30
- Levene P, LaForge F. 1941. On chondrosin. *J. Biol. Chem.* 140:267–77
- Levy P, Robert a, Picard J. 1988. Biosynthesis of glycosaminoglycans in the human colonic tumor cell line Caco-2: structural changes occurring with the morphological differentiation of the cells. *Biol. Cell.* 62(3):255–64
- Li J. 2010. Glucuronyl C5-Epimerase : An Enzyme Converting Glucuronic Acid to Iduronic Acid in Heparan Sulfate / Heparin Biosynthesis. *Prog. Molecular Biol. Transl. Sci.* 93:59–78
- Li W, Johnson DJD, Esmon CT, Huntington J a. 2004. Structure of the antithrombin-thrombin-heparin ternary complex reveals the antithrombotic mechanism of heparin. *Nat. Struct. Mol. Biol.* 11(9):857–62
- Lin H, Huber R, Schlessinger D, Morin PJ. 1999. Frequent Silencing of the GPC3 Gene in Ovarian Cancer Cell Lines. *Cancer Res.* 59:807–10

- Lin Y, Mahan K, Lathrop WF, Myles DG, Primakoff P. 1994. A hyaluronidase activity of the sperm plasma membrane protein PH-20 enables sperm to penetrate the cumulus cell layer surrounding the egg. *J. Cell Biol.* 125(5):1157–63
- Lindahl B, Eriksson L, Lindahl U. 1995. Structure of heparan sulphate from human brain , with special regard to Alzheimer ' s disease. *Biochem J.* 306:177–84
- Lindahl U, Bäckström G. 1972. Biosynthesis of L-iduronic acid in heparin : epimerization of D-glucuronic acid on the polymer level. *Biochem. Biophys. Res. Commun.* 46(2):985–91
- Lindahl U, Li J. 2009. *Interactions Between Heparan Sulfate and Proteins — Design and Functional Implications*, Vol. 276. Elsevier Inc. 1st ed.
- Linhardt RJ, Liu J. 2012. Synthetic heparin. *Curr Opin Pharmacol.* 12(2):217–19
- Linker A, Sampson P. 1960. The enzymic degradation of heparitin sulfate. *Biochim. Biophys. Acta.* 43:366–68
- Liu H, Zhang Z, Linhardt RJ. 2009. Lessons learned from the contamination of heparin. *Nat. Prod. Rep.* 26(3):313–21
- Liu J, Shriver Z, Pope RM, Thorp SC, Duncan MB, et al. 2002. Characterization of a heparan sulfate octasaccharide that binds to herpes simplex virus type 1 glycoprotein D. *J. Biol. Chem.* 277(36):33456–67
- Lohmander LS, Shinomura T, Hascall VC, Kimura JH. 1989. Xylosyl transfer to the core protein precursor of the rat chondrosarcoma proteoglycan. *J. Biol. Chem.* 264(31):18775–80
- López-Casillas F, Payne HM, Andres JL, Massagué J. 1994. Betaglycan can act as a dual modulator of TGF-beta access to signaling receptors: mapping of ligand binding and GAG attachment sites. *J. Cell Biol.* 124(4):557–68
- Lortat-Jacob H, Grimaud JA. 1991. Interferon-gamma C-terminal function : new working hypothesis . Heparan sulfate and heparin , new targets for IFN-gamma , protect , relax the cytokine and regulate its activity . *Cell Mol Biol.* 37(3):253–60
- Lortat-Jacob H, Grosdidier A, Imberty A. 2002. Structural diversity of heparan sulfate binding domains in chemokines. *Proc. Natl. Acad. Sci. U. S. A.* 99(3):1229–34
- Ludolph T, Paschke E, Glössl J, Kresse H. 1981. Degradation of keratan sulphate by beta-N-acetylhexosaminidases A and B. *Biochem. J.* 193(3):811–18
- Luo Y, Prestwich GD. 1999. Synthesis and Selective Cytotoxicity of a Hyaluronic Acid - Antitumor Bioconjugate. *Bioconjugate Chem.* 10(5):755–63
- Madine J, Clayton JC, Yates E a, Middleton D a. 2009. Exploiting a (13)C-labelled heparin analogue for in situ solid-state NMR investigations of peptide-glycan interactions within amyloid fibrils. *Org. Biomol. Chem.* 7(11):2414–20

- Madine J, Pandya MJ, Hicks MR, Rodger A, Yates EA, et al. 2012. Site-Specific Identification of an Abeta Fibril–Heparin Interaction Site by Using Solid-State NMR Spectroscopy. *Angew. Chemie*. 51:13140–43
- Mahalingam Y, Gallagher JT, Couchman JR. 2007. Cellular adhesion responses to the heparin-binding (HepII) domain of fibronectin require heparan sulfate with specific properties. *J. Biol. Chem*. 282(5):3221–30
- Mao Y, Schwarzbauer JE. 2005. Fibronectin fibrillogenesis, a cell-mediated matrix assembly process. *Matrix Biol*. 24(6):389–99
- Marcum J a. 2000. The origin of the dispute over the discovery of heparin. *J. Hist. Med. Allied Sci*. 55(1):37–66
- Marković-Housley Z, Miglierini G, Soldatova L, Rizkallah PJ, Müller U, Schirmer T. 2000. Crystal structure of hyaluronidase, a major allergen of bee venom. *Structure*. 8(10):1025–35
- Masamune H, Yosizawa Z, Maki M. 1951. Biochemical Studies on Carbohydrates : On Chondrosin and Chondroitinsulfuric Acid. *Tohoku J. Exp. Med*. 55(1):47–52
- Massagué J. 1985. Subunit Structure of a High-affinity Receptor for Type beta-transforming Frowth Factor. *J. Biol. Chem*. 260(11):7059–66
- Mathews M. 1958. Isomeric chondroitin sulphates. *Nature*. 181:184–85
- Matsuzaki H, Kobayashi H, Yagyu T, Wakahara K, Kondo T, et al. 2004. Bikunin Inhibits Lipopolysaccharide-Induced Tumor Necrosis Factor Alpha Induction in Macrophages. *Clin. Diagn. Lab. Immunol*. 11(6):1140–47
- Matter K. 2000. Epithelial polarity: sorting out the sorters. *Curr. Biol*. 10(1):R39–42
- Mayor S, Riezman H. 2004. Sorting GPI-anchored proteins. *Nat. Rev. Mol. Cell Biol*. 5(2):110–20
- Mccormick C, Duncan G, Goutsos KT, Tufaro F. 2000. The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate. *Proc. Natl. Acad. Sci. U. S. A*. 97(2):668–73
- Melnikova I. 2009. The anticoagulants market. *Drug Discov*. 8:353–54
- Merline R, Schaefer RM, Schaefer L. 2009. The matricellular functions of small leucine-rich proteoglycans (SLRPs). *J. Cell Commun. Signal*. 3(3-4):323–35
- Mertens G, Van der Schueren B, van den Berghe H, David G. 1996. Heparan sulfate expression in polarized epithelial cells: the apical sorting of glypican (GPI-anchored proteoglycan) is inversely related to its heparan sulfate content. *J. Cell Biol*. 132(3):487–97
- Meyer K, Davidson E, Linker A, Hoffman P. 1956. The acid mucopolysaccharides of connective tissue. *Biochim Biophys Acta*. 21:506–18

- Meyer K, Linker A, Eugene A, Weissmann B. 1953. THE MUCOPOLYSACCHARIDES OF BOVINE CORNEA. *J. Biol. Chem.* 205:611–16
- Meyer K, Palmer JW. 1934. The Polysaccharide of the Vitreous Humor. *J. Biol. Chem.* 107:629–34
- Meyer K, Smyth EM. 1937. On Glycoproteins: VI. The Preparation of Chondroitinsulfuric Acid. *J. Biol. Chem.* 119:507–10
- Michel G, Pojasek K, Li Y, Sulea T, Linhardt RJ, et al. 2004. The structure of chondroitin B lyase complexed with glycosaminoglycan oligosaccharides unravels a calcium-dependent catalytic machinery. *J. Biol. Chem.* 279(31):32882–96
- Middleton J, Neil S, Wintle J, Clark-Lewis I, Moore H, et al. 1997. Transcytosis and surface presentation of IL-8 by venular endothelial cells. *Cell.* 91(3):385–95
- Middleton J, Patterson AM, Gardner L, Schmutz C, Ashton BA. 2002. Leukocyte extravasation: chemokine transport and presentation by the endothelium. *Blood.* 100(12):3853–60
- Midura RJ, Calabro A, Yanagishita M, Hascall VC. 1995. Nonreducing end structures of chondroitin sulfate chains on aggrecan isolated from swarm rat chondrosarcoma cultures. *J. Biol. Chem.* 270(14):8009–15
- Miettinen HM, Edwards SN, Jalkanen M. 1994. Analysis of transport and targeting of syndecan-1: effect of cytoplasmic tail deletions. *Mol. Biol. Cell.* 5(12):1325–39
- Mikami T, Kitagawa H. 2013. Biosynthesis and function of chondroitin sulfate. *Biochim. Biophys. Acta.* 1830(10):4719–33
- Mikami T, Mizumoto S, Kago N, Kitagawa H, Sugahara K. 2003. Specificities of three distinct human chondroitin/dermatan N-acetylgalactosamine 4-O-sulfotransferases demonstrated using partially desulfated dermatan sulfate as an acceptor: implication of differential roles in dermatan sulfate biosynthesis. *J. Biol. Chem.* 278(38):36115–27
- Mikhailov D, Mayo KH, Vlahov IR, Toida T, Pervin A, Linhardt RJ. 1996. NMR solution conformation of heparin-derived tetrasaccharide. *Biochem J.* 318:93–102
- Misra S, Heldin P, Hascall VC, Karamanos NK, Skandalis SS, et al. 2011. HA/CD44 interactions as potential targets for cancer therapy. *FEBS J.* 278(9):1429–43
- Mohammadi M, Olsen SK, Ibrahim O a. 2005. Structural basis for fibroblast growth factor receptor activation. *Cytokine Growth Factor Rev.* 16(2):107–37
- Mondor I, Ugolini S, Sattentau QJ. 1998. Human immunodeficiency virus type 1 attachment to HeLa CD4 cells is CD4 independent and gp120 dependent and requires cell surface heparans. *J. Virol.* 72(5):3623–34
- Mongiati M, Taylor K, Otto J, Aho S, Uitto J, et al. 2000. The Protein Core of the Proteoglycan Perlecan Binds Specifically to Fibroblast Growth Factor-7. *J. Biol. Chem.* 275(10):7095–7100

- Moon LDF, Asher RA, Rhodes KE, Fawcett JW. 2001. Regeneration of CNS axons back to their target following treatment of adult rat brain with chondroitinase ABC. *Nat. Neurosci.* 4(5):465–66
- Morimoto-Tomita M, Uchimura K, Werb Z, Hemmerich S, Rosen SD. 2002. Cloning and characterization of two extracellular heparin-degrading endosulfatases in mice and humans. *J. Biol. Chem.* 277(51):49175–85
- Moses J, Oldberg A, Cheng F, Fransson L-A. 1997. Biosynthesis of the proteoglycan decorin. *Eur. J. Biochem.* 248:521–26
- Moses J, Oldberg a, Fransson L a. 1999. Initiation of galactosaminoglycan biosynthesis. Separate galactosylation and dephosphorylation pathways for phosphoxylosylated decorin protein and exogenous xyloside. *Eur. J. Biochem.* 260(3):879–84
- MRG. <http://mrg.net/News-and-Events/Press-Releases/Hyaluronic-Acid-Market-012813.aspx>
- Muir H. 1958. The nature of the link between protein and carbohydrate of a chondroitin sulphate complex from hyaline cartilage. *Biochem. J.* 69(2):195–204
- Müller A, Homey B, Soto H, Ge N, Catron D, et al. 2001. Involvement of chemokine receptors in breast cancer metastasis. *Nature.* 410:50–56
- Mulloy B, Forster MJ. 2000. Conformation and dynamics of heparin and heparan sulfate. *Glycobiology.* 10(11):1147–56
- Mulloy B, Forster MJ, Jones C, Davies DB. 1993. N.m.r. and molecular-modelling studies of the solution conformation of heparin. *Biochem. J.* 293:849–58
- Mulloy B, Forster MJ, Jones C, Drake AF, Johnson EA, Davies DB. 1994. The effect of variation of substitution on the solution conformation of heparin : a spectroscopic and molecular modelling study. *Carbohydr. Res.* 255:1–26
- Munesue S, Kusano Y, Oguri K, Itano N, Yoshitomi Y, et al. 2002. The role of syndecan-2 in regulation of actin-cytoskeletal organization of Lewis lung carcinoma-derived metastatic clones. *Biochem. J.* 363(Pt 2):201–9
- Murphy KJ, Mclay N, Pye DA. 2008. Structural Studies of Heparan Sulfate Hexasaccharides : New Insights into Iduronate Conformational Behavior. *J Am Chem Soc.* 130(37):12435–44
- Mythreya K, Blobel GC. 2009. Proteoglycan signaling co-receptors: roles in cell adhesion, migration and invasion. *Cell. Signal.* 21(11):1548–58
- Nadanaka S, Kitagawa H. 2008. Heparan sulphate biosynthesis and disease. *J. Biochem.* 144(1):7–14
- Nakajima M, Irimura T, Nicolson GL. 1988. Heparanases and Tumor Metastasis. *J. Cell. Biochem.* 36:157–67

- Nakazawa K, Takahashi I, Yamamoto Y. 1998. Glycosyltransferase and sulfotransferase activities in chick corneal stromal cells before and after in vitro culture. *Arch. Biochem. Biophys.* 359(2):269–82
- Ng L, Grodzinsky AJ, Patwari P, Sandy J, Plaas A, Ortiz C. 2003. Individual cartilage aggrecan macromolecules and their constituent glycosaminoglycans visualized via atomic force microscopy. *J. Struct. Biol.* 143(3):242–57
- Nieduszynski I a, Atkins EDT. 1973. Conformation of the Mucopolysaccharides. X-Ray Fibre Diffraction of Heparin. *Biochem J.* 135:729–33
- Nili N, Cheema AN, Giordano FJ, Barolet AW, Babaei S, et al. 2003. Decorin inhibition of PDGF-stimulated vascular smooth muscle cell function: potential mechanism for inhibition of intimal hyperplasia after balloon angioplasty. *Am. J. Pathol.* 163(3):869–78
- Norgard-Sumnicht K, Varki A. 1995. Endothelial Heparan Sulfate Proteoglycans That Bind to L-Selectin Have Glucosamine Residues with Unsubstituted Amino Groups. *J. Biol. Chem.* 270(20):12012–24
- Noulas A V, Skandalis SS, Feretis E, Theocharis D a, Karamanos NK. 2004. Variations in content and structure of glycosaminoglycans of the vitreous gel from different mammalian species. *Biomed. Chromatogr.* 18(7):457–61
- Oh E-S, Couchman JR. 2004. Syndecans-2 and-4; Close Cousins, but not Identical Twins. *Mol. Cells.* 17(2):181–87
- Ohtake-Niimi S, Kondo S, Ito T, Kakehi S, Ohta T, et al. 2010. Mice deficient in N-acetylgalactosamine 4-sulfate 6-o-sulfotransferase are unable to synthesize chondroitin/dermatan sulfate containing N-acetylgalactosamine 4,6-bissulfate residues and exhibit decreased protease activity in bone marrow-derived mast cells. *J. Biol. Chem.* 285(27):20793–805
- Okajima T, Yoshida K, Kondo T, Furukawa K. 1999. Human Homolog of *Caenorhabditis elegans* sqv-3 Gene Is Galactosyltransferase I Involved in the Biosynthesis of the Glycosaminoglycan-Protein Linkage Region of Proteoglycans. *J. Biol. Chem.* 274(33):22915–18
- Okuyama E, Suzuki A, Murata M, Ando Y, Kato I, et al. 2013. Molecular mechanisms of syndecan-4 upregulation by TNF- α in the endothelium-like EAhy926 cells. *J. Biochem.* 154(1):41–50
- Ori A, Wilkinson MC, Fernig DG. 2008. The heparanome and regulation of cell function : structures, functions and challenges. *Front. Biosci.* 13:4309–38
- Pacheco B, Malmström A, Maccarana M. 2009. Two dermatan sulfate epimerases form iduronic acid domains in dermatan sulfate. *J. Biol. Chem.* 284(15):9788–95
- Pankov R, Yamada K. 2002. Fibronectin at a glance. *J. Cell Sci.* 115(20):3861–63
- Parish CR. 2006. The role of heparan sulphate in inflammation. *Nat. Rev. Immunol.* 6(9):633–43

- Park PW, Pier GB, Hinkes MT, Bernfield M. 2001. Exploitation of syndecan-1 shedding by *Pseudomonas aeruginosa* enhances virulence. *Nature*. 411(6833):98–102
- Partenheimer A, Schwarz K, Wrocklage C, Koisch E, Kresse H. 1995. Proteoglycan Form of Colony-Stimulating Factor-1 (Proteoglycan-100): Stimulation of Activity by Glycosaminoglycan Removal and Proteolytic Processing. *J. Immunol.*, pp. 5557–65
- Pegeot M, Sadir R, Eriksson I, Kjellén L, Simorre J, et al. 2014. Profiling sulfation/epimerization pattern of full-length heparan sulfate by NMR following cell culture ¹³C-glucose metabolic labeling. *Glycobiology*
- Pellegrini L, Burke DF, von Delft F, Mulloy B, Blundell TL. 2000. Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature*. 407(6807):1029–34
- Pethe K, Alonso S, Biet F, Delogu G, Brennan MJ, et al. 2001. The heparin-binding haemagglutinin of *M. tuberculosis* is required for extrapulmonary dissemination. *Nature*. 412(6843):190–94
- Petitou M, Casu B, Lindahl U. 2003. 1976-1983, a Critical Period in the History of Heparin: the Discovery of the Antithrombin Binding Site. *Biochimie*. 85(1-2):83–89
- Philipson LH, Westley J, Schwartz NB. 1985. Effect of hyaluronidase treatment of intact cells on hyaluronate synthetase activity. *Biochemistry*. 24(27):7899–7906
- Pienimäki JP, Rilla K, Fulop C, Sironen RK, Karvinen S, et al. 2001. Epidermal growth factor activates hyaluronan synthase 2 in epidermal keratinocytes and increases pericellular and intracellular hyaluronan. *J. Biol. Chem*. 276(23):20428–35
- Pinhal M a, Smith B, Olson S, Aikawa J, Kimata K, Esko JD. 2001. Enzyme interactions in heparan sulfate biosynthesis: uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo. *Proc. Natl. Acad. Sci. U. S. A*. 98(23):12984–89
- Plaas a H, West L a, Midura RJ. 2001. Keratan sulfate disaccharide composition determined by FACE analysis of keratanase II and endo-beta-galactosidase digestion products. *Glycobiology*. 11(10):779–90
- Poluri KM, Joseph PRB, Sawant K V, Rajarathnam K. 2013. Molecular Basis of Glycosaminoglycan Heparin Binding to Chemokine CXCL1 dimer. *J. Biol. Chem.*, pp. 1–22
- Pomin VH. 2013. NMR Chemical Shifts in Structural Biology of Glycosaminoglycans. *Anal. Chem*. 86(1):65–94
- Pomin VH. 2014. Biological findings from the recent NMR-based studies of glycosaminoglycan–protein interactions. *Glycobiology*. 0(0):1–13
- Pomin VH, Sharp JS, Li X, Wang L, Prestegard JH. 2010. Characterization of glycosaminoglycans by ¹⁵N NMR spectroscopy and in vivo isotopic labeling. *Anal. Chem*. 82(10):4078–88

- Ponta H, Sherman L, Herrlich P a. 2003. CD44: from adhesion molecules to signalling regulators. *Nat. Rev. Mol. Cell Biol.* 4(1):33–45
- Prehm P. 1983. Synthesis of hyaluronate in differentiated teratocarcinoma cells: mechanism of chain growth. *Biochem J.* 211:191–98
- Prehm P. 1984. Hyaluronate is synthesized at plasma membranes. *Biochem. J.* 220(2):597–600
- Presto J, Thuveson M, Carlsson P, Busse M, Wilén M, et al. 2008. Heparan sulfate biosynthesis enzymes EXT1 and EXT2 affect NDST1 expression and heparan sulfate sulfation. *Proc. Natl. Acad. Sci. U. S. A.* 105(12):4751–56
- Printz JO, Lee JJ, Knesek M, Urquhart AG. 2013. Conflict of Interest in the Assessment of Hyaluronic Acid Injections for Osteoarthritis of the Knee: An Updated Systematic Review. *J. Arthroplasty.* 28(8):30–33.e1
- Prydz K, Dalen KT. 2000. Synthesis and sorting of proteoglycans. *J. Cell Sci.* 113 Pt 2:193–205
- Prydz K, Dick G, Tveit H. 2008. How many ways through the Golgi maze? *Traffic.* 9(3):299–304
- Pye D a, Vivès RR, Hyde P, Gallagher JT. 2000. Regulation of FGF-1 mitogenic activity by heparan sulfate oligosaccharides is dependent on specific structural features: differential requirements for the modulation of FGF-1 and FGF-2. *Glycobiology.* 10(11):1183–92
- Pye D a., Vives RR, Turnbull JE, Hyde P, Gallagher JT. 1998. Heparan Sulfate Oligosaccharides Require 6-O-Sulfation for Promotion of Basic Fibroblast Growth Factor Mitogenic Activity. *J. Biol. Chem.* 273(36):22936–42
- Rapraeger a C, Krufka a, Olwin BB. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science.* 252(5013):1705–8
- Rauch U, Feng K, Zhou X. 2001. Neurocan : a brain chondroitin sulfate. *Cell. Mol. Life Sci.* 58(12-13):1842–56
- Rees MD, McNiven TN, Davies MJ. 2007. Degradation of extracellular matrix and its components by hypobromous acid. *Biochem. J.* 401(2):587–96
- Reisher SR, Hughesti TE, Ordovast JM, Schaefer EJ, Feinstein SI. 1993. Increased expression of apolipoprotein genes accompanies differentiation in the intestinal cell line Caco-2. *Proc. Natl. Acad. Sci. U. S. A.* 90:5757–61
- Renault M, Cukkemane A, Baldus M. 2010. Solid-State NMR Spectroscopy on Complex Biomolecules. *Angew. Chemie.* 49:8346–57
- Renault M, Tommassen-van Boxtel R, Bos MP, Post JA, Tommassen J, Baldus M. 2012. Cellular solid-state nuclear magnetic resonance spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 109(13):4863–68

- Rhomberg a J, Ernst S, Sasisekharan R, Biemann K. 1998. Mass spectrometric and capillary electrophoretic investigation of the enzymatic degradation of heparin-like glycosaminoglycans. *Proc. Natl. Acad. Sci. U. S. A.* 95(8):4176–81
- Ricard-Blum S. 2011. The collagen family. *Cold Spring Harb. Perspect. Biol.* 3(1):a004978
- Ricard-Blum S, Féraud O, Lortat-Jacob H, Rencurosi A, Fukai N, et al. 2004. Characterization of endostatin binding to heparin and heparan sulfate by surface plasmon resonance and molecular modeling: role of divalent cations. *J. Biol. Chem.* 279(4):2927–36
- Ringvall M, Ledin J, Holmborn K, van Kuppevelt T, Ellin F, et al. 2000. Defective heparan sulfate biosynthesis and neonatal lethality in mice lacking N-deacetylase/N-sulfotransferase-1. *J. Biol. Chem.* 275(34):25926–30
- Ripellino J a, Bailo M, Margolis RU, Margolis RK. 1988. Light and electron microscopic studies on the localization of hyaluronic acid in developing rat cerebellum. *J. Cell Biol.* 106(3):845–55
- Roden L, Armand G. 1966. Structure of the Chondroitin 4-Sulfate-Protein Linkage Region : ISOLATION AND CHARACTERIZATION OF THE DISACCHARIDE 3-O-BETA-D-GLUCURONOSYL-D-GALACTOSE. *J. Biol. Chem.* 241(1):65–70
- Roden L, Smith R. 1966. Structure of the Neutral Trisaccharide of the Chondroitin 4-Sulfate-Protein Linkage Region. *J. Biol. Chem.* 241(24):5949–54
- Rong J, Habuchi H, Kimata K, Lindahl U, Kusche-Gullberg M. 2001. Substrate specificity of the heparan sulfate hexuronic acid 2-O-sulfotransferase. *Biochemistry.* 40(18):5548–55
- Rot A. 1992. Endothelial cell binding of NAP-1 / IL-8 : role in neutrophil emigration. *Immunol. Today.* 13(8):291–94
- Rousset M. 1986. The human colon carcinoma cell lines HT-29 and Caco-2: two in vitro models for the study of intestinal differentiation. *Biochimie.* 68(9):1035–40
- Rudd TR, Guimond SE, Skidmore MA, Duchesne L, Guerrini M, et al. 2007. Influence of substitution pattern and cation binding on conformation and activity in heparin derivatives. *Glycobiology.* 17(9):983–93
- Ruoslahti E. 1989. Proteoglycans in Cell Regulation. *J. Biol. Chem.* 264(23):13369–72
- Sadir R, Imberty A, Baleux F, Lortat-Jacob H. 2004. Heparan sulfate/heparin oligosaccharides protect stromal cell-derived factor-1 (SDF-1)/CXCL12 against proteolysis induced by CD26/dipeptidyl peptidase IV. *J. Biol. Chem.* 279(42):43854–60
- Saesen E, Sarrazin S, Laguri C, Sadir R, Maurin D, et al. 2013. Insights into the Mechanism by Which Interferon - γ Basic Amino Acid Clusters Mediate Protein Binding to Heparan Sulfate. *J Am Chem Soc.* 135:9384–90
- Saksela O, Moscatelli D, Sommer a, Rifkin DB. 1988. Endothelial cell-derived heparan sulfate binds basic fibroblast growth factor and protects it from proteolytic degradation. *J. Cell Biol.* 107(2):743–51

- Salmivirta M, Safaiyan F, Prydz K, Andresen MS, Aryan M, Kolset SO. 1998. Differentiation-associated modulation of heparan sulfate structure and function in CaCo-2 colon carcinoma cells. *Glycobiology*. 8(10):1029–36
- Sandy J, Flannery C, Boynton R, Neames P. 1990. Isolation and Characterization of Disulfide-bonded Peptides from the Three Globular Domains of Aggregating Cartilage Proteoglycan. *J. Biol. Chem.* 265(34):21108–13
- Sapay N, Cabannes E, Petitou M, Imberty A. 2011. Molecular modeling of the interaction between heparan sulfate and cellular growth factors: bringing pieces together. *Glycobiology*. 21(9):1181–93
- Sarrazin S, Lamanna WC, Esko JD. 2011. Heparan sulfate proteoglycans. *Cold Spring Harb. Perspect. Biol.* 3(7):1–33
- Saunders S, Paine-Saunders S, Lander a D. 1997. Expression of the cell surface proteoglycan glypican-5 is developmentally regulated in kidney, limb, and brain. *Dev. Biol.* 190(1):78–93
- Schaefer L, Babelova A, Kiss E, Hausser H, Baliova M, et al. 2005. The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages. *J. Clin. Invest.* 115(8):2223–33
- Scheidt H a, Schibur S, Magalhães A, de Azevedo ER, Bonagamba TJ, et al. 2010. The mobility of chondroitin sulfate in articular and artificial cartilage characterized by ¹³C magic-angle spinning NMR spectroscopy. *Biopolymers*. 93(6):520–32
- Scheiffele P, Peränen J, Simons K. 1995. N-glycans as apical sorting signals in epithelial cells. *Nature*. 378:96–98
- Schlessinger J, Plotnikov a N, Ibrahimi O a, Eliseenkova a V, Yeh BK, et al. 2000. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell*. 6(3):743–50
- Schmiedeberg O. 1891. Über die chemische Zusammensetzung des Knorpels. *Arch Exptl Pathol Pharmacol*. 28:354–404
- Schulz T, Schumacher U, Prehm P. 2007. Hyaluronan export by the ABC transporter MRP5 and its modulation by intracellular cGMP. *J. Biol. Chem.* 282(29):20999–4
- Schwarz K, Breuer B, Kresse H. 1990. Biosynthesis and properties of a further member of the small chondroitin/dermatan sulfate proteoglycan family. *J. Biol. Chem.* 265(35):22023–28
- Scott JE, Heatley F. 2002. Biological properties of hyaluronan in aqueous solution are controlled and sequestered by reversible tertiary structures, defined by NMR spectroscopy. *Biomacromolecules*. 3(3):547–53
- Scott PG, Dodd CM, Bergmann EM, Sheehan JK, Bishop PN. 2006. Crystal structure of the biglycan dimer and evidence that dimerization is essential for folding and stability of class I small leucine-rich repeat proteoglycans. *J. Biol. Chem.* 281(19):13324–32

- Sedita J, Izvolsky K, Cardoso W V. 2004. Differential expression of heparan sulfate 6-O-sulfotransferase isoforms in the mouse embryo suggests distinctive roles during organogenesis. *Dev. Dyn.* 231(4):782–94
- Seffouh A, Milz F, Przybylski C, Laguri C, Oosterhof A, et al. 2013. HSulf sulfatases catalyze processive and oriented 6-O-desulfation of heparan sulfate that differentially regulates fibroblast growth factor activity. *FASEB J.* 27(6):2431–39
- Segarini PR, Seyedin SM. 1988. The high molecular weight receptor to transforming growth factor-beta contains glycosaminoglycan chains. *J. Biol. Chem.* 263(17):8366–70
- Seko A, Yamashita K. 2004. β 1,3-N-Acetylglucosaminyltransferase-7 (β 3Gn-T7) acts efficiently on keratan sulfate-related glycans. *FEBS Lett.* 556(1-3):216–20
- Seppinen L, Pihlajaniemi T. 2011. The multiple functions of collagen XVIII in development and disease. *Matrix Biol.* 30:83–92
- Shatton J, Schubert M. 1954. Isolation of a Mucoprotein from Cartilage. *J. Biol. Chem.* 211:565–73
- Shaya D, Zhao W, Garron M-L, Xiao Z, Cui Q, et al. 2010. Catalytic mechanism of heparinase II investigated by site-directed mutagenesis and the crystal structure with its substrate. *J. Biol. Chem.* 285(26):20051–61
- Shi X, Zaia J. 2009. Organ-specific Heparan Sulfate Structural Phenotypes. *J. Biol. Chem.* 284(18):11806–14
- Shively JE, Conrad HE. 1976. Formation of anhydrosugars in the chemical depolymerization of heparin. *Biochemistry.* 15(18):3932–42
- Shukla D, Liu J, Blaiklock P, Shworak NW, Bai X, et al. 1999. A Novel Role for 3- O -Sulfated Heparan Sulfate in Herpes Simplex Virus 1 Entry. *Ce.* 99:13–22
- Shworak NW, Liu J, Petros LM, Zhang L, Kobayashi M, et al. 1999. Multiple Isoforms of Heparan Sulfate D-Glucosaminyl 3-O-Sulfotransferase: ISOLATION, CHARACTERIZATION, AND EXPRESSION OF HUMAN cDNAs AND IDENTIFICATION OF DISTINCT GENOMIC LOCI. *J. Biol. Chem.* 274(8):5170–84
- Sibille N, Sillen A, Leroy A, Wieruszeski J, Mulloy B, et al. 2006. Structural Impact of Heparin Binding to Full-Length Tau As Studied by NMR. *Biochemistry.* 45:12560–72
- Sillen A, Leroy A, Wieruszeski J, Loyens A, Beauvillain C, et al. 2005a. Regions of Tau Implicated in the Paired Helical Fragment Core as Defined by NMR. *Chembiochem.* 6:1849–56
- Sillen A, Wieruszeski J, Leroy A, Younes A Ben, Landrieu I, Lippens G. 2005b. High-Resolution Magic Angle Spinning NMR of the Neuronal Tau Protein Integrated in Alzheimer's-Like Paired Helical Fragments. *J Am Chem Soc.* 127:10138–39

- Skidmore M a, Guimond SE, Dumax-Vorzet AF, Yates E a, Turnbull JE. 2010. Disaccharide compositional analysis of heparan sulfate and heparin polysaccharides using UV or high-sensitivity fluorescence (BODIPY) detection. *Nat. Protoc.* 5(12):1983–92
- Song C, Hu K, Joo C, Swager TM, Griffin RG. 2006. TOTAPOL : A Biradical Polarizing Agent for Dynamic Nuclear Polarization Experiments in Aqueous Media. *J Am Chem Soc.* 128(35):11385–90
- Song HH, Shi W, Filmus J. 1997. OCI-5/Rat Glypican-3 Binds to Fibroblast Growth Factor-2 but Not to Insulin-like Growth Factor-2. *J. Biol. Chem.* 272(12):7574–77
- Song HH, Shi W, Xiang Y-Y, Filmus J. 2005. The loss of glypican-3 induces alterations in Wnt signaling. *J. Biol. Chem.* 280(3):2116–25
- Spear PG. 2004. Herpes simplex virus : receptors and ligands for cell entry. *Cell. Microbiol.* 6(5):401–10
- Spicer AP, Seldin MF, Olsen AS, Brown N, Wells DE, et al. 1997. Chromosomal localization of the human and mouse hyaluronan synthase genes. *Genomics.* 41(3):493–97
- Staples GO, Shi X, Zaia J. 2011. Glycomics analysis of mammalian heparan sulfates modified by the human extracellular sulfatase HSulf2. *PLoS One.* 6(2):e16689
- Stern R. 2003. Devising a pathway for hyaluronan catabolism: are we there yet? *Glycobiology.* 13(12):105R–115R
- Stevenson JL, Choi SH, Varki A. 2005. Differential Metastasis Inhibition by Clinically Relevant Levels of Heparins — Correlation with Selectin Inhibition , Not Antithrombotic Activity. *Clin Cancer Res.* 11:7003–11
- Strand ME, Herum KM, Rana Z a, Skrbic B, Askevold ET, et al. 2013. Innate immune signaling induces expression and shedding of the heparan sulfate proteoglycan syndecan-4 in cardiac fibroblasts and myocytes, affecting inflammation in the pressure-overloaded heart. *FEBS J.* 280(10):2228–47
- Suzuki M. 1939. BIOCHEMICAL STUDIES ON CARBOHYDRATES. L. Prosthetic Group of Corneamucoid. *J. Biochem.* 30(2):185–91
- Svensson L, Heinegard D, Oldberg A. 1995. Decorin-binding Sites for Collagen Type I Are Mainly Located in Leucine-rich Repeats 4-5. *J. Biol. Chem.* 270(35):20712–16
- Szatmári T, Dobra K. 2013. The role of syndecan-1 in cellular signaling and its effects on heparan sulfate biosynthesis in mesenchymal tumors. *Front. Oncol.* 3:1–11
- Tagalakakis V, Blostein M, Robinson-Cohen C, Kahn SR. 2007. The effect of anticoagulants on cancer risk and survival : Systematic review. *Cancer Treat. Rev.* 33:358–68
- Tai G-H, Huckerby TN, Nieduszynski IA. 1996. Multiple Non-reducing Chain Termini Isolated from Bovine Corneal Keratan Sulfates. *J. Biol. Chem.* 271(38):23535–46

- Takahashi H, Ayala I, Bardet M, De Paëpe G, Simorre J, Hediger S. 2013. Solid-State NMR on Bacterial Cells: Selective Cell Wall Signal Enhancement and Resolution Improvement using Dynamic Nuclear Polarization. *J Am Chem Soc.* 135:5105–10
- Takegawa Y, Araki K, Fujitani N, Furukawa J, Sugiyama H, et al. 2011. Simultaneous analysis of heparan sulfate, chondroitin/dermatan sulfates, and hyaluronan disaccharides by glycoblotting-assisted sample preparation followed by single-step zwitter-ionic-hydrophilic interaction chromatography. *Anal. Chem.* 83(24):9443–49
- Tanaka Y, Adams DH, Hubscher S, Hirano H, Siebenlist U, Shaw S. 1993. T-cell adhesion induced by proteoglycan-immobilized cytokine MIP-1 beta. *Nature.* 361:79–82
- Taylor RL, Shively J., Conrad HE, Cifonelli A. 1973. Uronic Acid Composition of Heparins and Heparan Sulfates. *Biochemistry.* 12(19):3633–37
- Tollefsen DM. 2010. Vascular Dermatan Sulfate and Heparin Cofactor II. *Prog. Mol. Biol. Transl. Sci.* 93(10):351–72
- Trescony P V, Oegama TRJ, Farnam BJ, Deloria LB. 1989. Analysis of heparan sulfate from the Engelbreth-Holm-Swarm (EHS) tumor. *Connect Tissue Res.* 19(2-4):219–42
- Trowbridge JM, Gallo RL. 2002. Dermatan sulfate: new functions from an old glycosaminoglycan. *Glycobiology.* 12(9):117R–25R
- Tufvesson E, Westergren-Thorsson G. 2002. Tumour necrosis factor-alpha interacts with biglycan and decorin. *FEBS Lett.* 530:124–28
- Turley RS, Finger EC, Hempel N, How T, Fields T a, Blobe GC. 2007. The type III transforming growth factor-beta receptor as a novel tumor suppressor gene in prostate cancer. *Cancer Res.* 67(3):1090–98
- Tvaroška I, Hricovini M, Petrakova E. 1989. An attempt to derive a new Karplus-type equation of vicinal proton-carbon coupling constants for C-O-C-H segments of bonded atoms. *Carbohydr. Res.* 189:359–62
- Tvaroška I, Pérez S. 1986. Conformational-energy calculations for oligosaccharides: a comparison of methods and a strategy of calculation. *Carbohydr. Res.* 149(2):389–410
- Uchimura K, Kadomatsu K, Fan QW, Muramatsu H, Kurosawa N, et al. 1998. Mouse chondroitin 6-sulfotransferase: molecular cloning, characterization and chromosomal mapping. *Glycobiology.* 8(5):489–96
- Ueno M, Yamada S, Zako M, Bernfield M, Sugahara K. 2001. Structural characterization of heparan sulfate and chondroitin sulfate of syndecan-1 purified from normal murine mammary gland epithelial cells. Common phosphorylation of xylose and differential sulfation of galactose in the protein linkage region tetrasa. *J. Biol. Chem.* 276(31):29134–40
- Uhlin-Hansen L, Eskeland T, Kolset SO. 1989. Modulation of the expression of chondroitin sulfate proteoglycan in stimulated human monocytes. *J. Biol. Chem.* 264(25):14916–22

- Ujita M, Misra a K, McAuliffe J, Hindsgaul O, Fukuda M. 2000. Poly-N-acetyllactosamine extension in N-glycans and core 2- and core 4-branched O-glycans is differentially controlled by i-extension enzyme and different members of the beta 1,4-galactosyltransferase gene family. *J. Biol. Chem.* 275(21):15868–75
- Uyama T, Kitagawa H, Tamura Ji J, Sugahara K. 2002. Molecular cloning and expression of human chondroitin N-acetylgalactosaminyltransferase: the key enzyme for chain initiation and elongation of chondroitin/dermatan sulfate on the protein linkage region tetrasaccharide shared by heparin/heparan sulfate. *J. Biol. Chem.* 277(11):8841–46
- Uyama T, Kitagawa H, Tanaka J, Tamura J, Ogawa T, Sugahara K. 2003. Molecular cloning and expression of a second chondroitin N-acetylgalactosaminyltransferase involved in the initiation and elongation of chondroitin/dermatan sulfate. *J. Biol. Chem.* 278(5):3072–78
- Van Zijl F Van, Krupitza G, Mikulits W. 2011. Initial steps of metastasis : Cell invasion and endothelial transmigration. *Mutat. Res.* 728:23–34
- Vanpouille C, Deligny A, Delehedde M, Denys A, Melchior A, et al. 2007. The heparin/heparan sulfate sequence that interacts with cyclophilin B contains a 3-O-sulfated N-unsubstituted glucosamine residue. *J. Biol. Chem.* 282(33):24416–29
- Velasco-Loyden G, Arribas J, López-Casillas F. 2004. The shedding of betaglycan is regulated by pervanadate and mediated by membrane type matrix metalloprotease-1. *J. Biol. Chem.* 279(9):7721–33
- Venkataraman G. 1999. Sequencing Complex Polysaccharides. *Science (80-.)*. 286(5439):537–42
- Vertel BM, Walters LM, Flay N, Kearns a E, Schwartz NB. 1993. Xylosylation is an endoplasmic reticulum to Golgi event. *J. Biol. Chem.* 268(15):11105–12
- Veugelers M, De Cat B, Ceulemans H, Bruystens a.-M, Coomans C, et al. 1999. Glypican-6, a New Member of the Glypican Family of Cell Surface Heparan Sulfate Proteoglycans. *J. Biol. Chem.* 274(38):26968–77
- Vivès RR, Pye D a, Salmivirta M, Hopwood JJ, Lindahl U, Gallagher JT. 1999. Sequence analysis of heparan sulphate and heparin oligosaccharides. *Biochem. J.* 339 (Pt 3:767–73
- Vlodavsky I, Bashkin P, Ishai-Michaeli R, Chajek-Shaul T, Bar-Shavit R, et al. 1991. Sequestration and Release of Basic Fibroblast Growth Factor. *Ann. N. Y. Acad. Sci.* 638:207–20
- Vlodavsky I, Friedmann Y, Elkin M, Aingorn H, Atzmon R, et al. 1999. Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis. *Nat. Med.* 5(7):793–802
- Vlodavsky I, Goldshmidt O, Zcharia E, Atzmon R, Rangini-Guatta Z, et al. 2002. Mammalian heparanase: involvement in cancer metastasis, angiogenesis and normal development. *Semin. Cancer Biol.* 12(2):121–29

- Vuong TT, Prydz K, Tveit H. 2006. Differences in the apical and basolateral pathways for glycosaminoglycan biosynthesis in Madin-Darby canine kidney cells. *Glycobiology*. 16(4):326–32
- Wang XF, Lin HY, Ng-Eaton E, Downward J, Lodish HF, Weinberg R a. 1991. Expression cloning and characterization of the TGF-beta type III receptor. *Cell*. 67(4):797–805
- Weissmann B, Meyer K. 1954. The Structure of Hyalobiuronic Acid and of Hyaluronic Acid from Umbilical Cord. . 76:1753–57
- Wenstrup RJ, Florer JB, Brunskill EW, Bell SM, Chervoneva I, Birk DE. 2004. Type V collagen controls the initiation of collagen fibril assembly. *J. Biol. Chem*. 279(51):53331–37
- Wernersson S, Pejler G. 2014. Mast cell secretory granules: armed for battle. *Nat. Rev. Immunol*. 14(7):478–94
- Wessels MR, Moses AE, Goldberg JB, Dicesare TJ. 1991. Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci. *Proc. Natl. Acad. Sci. U. S. A*. 88:8317–21
- Winter WT, Arnott S. 1977. Hyaluronic acid: the role of divalent cations in conformation and packing. *J. Mol. Biol*. 117(3):761–84
- Winter WT, Arnott S, Isaac DH, Atkins EDT. 1978. Chondroitin 4-Sulfate : The Structure of a Sulfated Glycosaminoglycan. *J Mol Biol*. 125:1–19
- Wise SG, Weiss AS. 2009. Tropoelastin. *Int. J. Biochem. Cell Biol*. 41(3):494–97
- Wishart D, Bigam C, Yao J, Abildgaard F, Dyson HJ, et al. 1995. ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*. 6(2):8589602
- Wobig M, Bach G, Beks P, Dickhut A, Runzheimer J, et al. 1999. The Role of Elastoviscosity in the Efficacy of Viscosupplementation for Osteoarthritis of the Knee : A Comparison of Hylan G-F 20 and a Lower-Molecular- Weight Hyaluronan. *Clin. Ther*. 21(9):1549–62
- Woessner JF. 1994. The family of matrix metalloproteinases. *Ann. N. Y. Acad. Sci*. 732:11–21
- Wraith JE, Scarpa M, Beck M, Bodamer OA, De Meirleir L, et al. 2008. Mucopolysaccharidosis type II (Hunter syndrome): a clinical review and recommendations for treatment in the era of enzyme replacement therapy. *Eur J Pediatr*. 167:267–77
- Wu J-J, Weis MA, Kim LS, Eyre DR. 2010. Type III collagen, a fibril network modifier in articular cartilage. *J. Biol. Chem*. 285(24):18537–44
- Wüthrich K. 1986. *NMR of Proteins and Nucleic Acids*. New York
- Xu D, Tiwari V, Xia G, Clement C, Shukla D, Liu J. 2005. Characterization of heparan sulphate 3-O-sulphotransferase isoform 6 and its role in assisting the entry of herpes simplex virus type 1. *Biochem. J*. 385(Pt 2):451–59

- Yamada S, Okada Y, Ueno M, Iwata S, Deepa SS, et al. 2002. Determination of the glycosaminoglycan-protein linkage region oligosaccharide structures of proteoglycans from *Drosophila melanogaster* and *Caenorhabditis elegans*. *J. Biol. Chem.* 277(35):31877–86
- Yamada T, Kawasaki T. 2005. Microbial Synthesis of Hyaluronan and Chitin: New Approaches. *J Biosci Bioeng.* 99(6):521–28
- Yamada T, Ohtake S, Sato M, Habuchi O. 2004. Chondroitin 4-sulphotransferase-1 and chondroitin 6-sulphotransferase-1 are affected differently by uronic acid residues neighbouring the acceptor GalNAc residues. *Biochem. J.* 384(Pt 3):567–75
- Yamaguchi Y. 1996. Brevican: a major proteoglycan in adult brain. *Perspect. Dev. Neurobiol.* 3(4):307–17
- Yamamoto Y, Takahashi I, Ogata N, Nakazawa K. 2001. Purification and characterization of N-acetylglucosaminyl sulfotransferase from chick corneas. *Arch. Biochem. Biophys.* 392(1):87–92
- Yanagishita M, Hascall VC. 1992. Cell Surface Heparan Sulfate - Proteoglycans. *J. Biol. Chem.* 267(14):9451–54
- Yates E a, Santini F, Guerrini M, Naggi a, Torri G, Casu B. 1996. ¹H and ¹³C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives. *Carbohydr. Res.* 294:15–27
- Yayon a, Klagsbrun M, Esko JD, Leder P, Ornitz DM. 1991. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell.* 64(4):841–48
- Yu F, Wolff JJ, Amster IJ, Prestegard JH. 2007. Conformational Preferences of Chondroitin Sulfate Oligomers Using Partially Oriented NMR Spectroscopy of ¹³C-Labeled Acetyl Groups. *J Am Chem Soc.* 129:13288–97
- Zaia J, Li X-Q, Chan S-Y, Costello CE. 2003. Tandem mass spectrometric strategies for determination of sulfation positions and uronic acid epimerization in chondroitin sulfate oligosaccharides. *J. Am. Soc. Mass Spectrom.* 14(11):1270–81
- Zetser A, Bashenko Y, Edovitsky E, Levy-Adam F, Vlodavsky I, Ilan N. 2006. Heparanase Induces Vascular Endothelial Growth Factor Expression: Correlation with p38 Phosphorylation Levels and Src Activation. *Cancer Res.* 66(3):1455–63
- Zhang L, Beeler DL, Lawrence R, Lech M, Liu J, et al. 2001. 6-O-sulfotransferase-1 represents a critical enzyme in the anticoagulant heparan sulfate biosynthetic pathway. *J. Biol. Chem.* 276(45):42311–21
- Zhang Z, McCallum S a, Xie J, Nieto L, Corzana F, et al. 2008. Solution structures of chemoenzymatically synthesized heparin and its precursors. *J. Am. Chem. Soc.* 130(39):12998–7

Ziarek JJ, Veldkamp CT, Zhang F, Murray NJ, Kartz A, et al. 2013. Heparin Oligosaccharides Inhibit Chemokine (CXC Motif) Ligand 12 (CXCL12) Cardioprotection by Binding Orthogonal to the Dimerization Interface , Promoting Oligomerization , and Competing with the Chemokine (CXC Motif) Receptor 4 N Terminus. *J. Biol. Chem.* 288(1):737–46

Zöller M. 2011. CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? *Nat. Rev. Cancer.* 11(4):254–67

Résumé

Les glycosaminoglycanes (GAG) forment une famille de polysaccharides linéaires retrouvés dans tous les tissus, au niveau des matrices extracellulaires et des surfaces cellulaires. Les héparanes sulfates (HS) sont des membres importants de cette famille et sont liés à une protéine dite cœur pour former ensemble le protéoglycane (PG). Selon le tissu et la nature de la protéine cœur, les HS, composés d'unités disaccharidiques de *N*-acétylglucosamine (GlcNAc) et d'acide glucuronique (GlcA) [-4GlcA β 1-4GlcNAc α 1-] vont subir de nombreuses modifications. En effet, les HS sont modifiés par différentes sulfatations au niveau des deux oses et une épimérisation de l'acide glucuronique en acide iduronique (IdoA). Les différentes structures saccharidiques élaborées vont pouvoir interagir avec une très grande quantité de protéines et jouer des rôles divers dans l'inflammation, la prolifération cellulaire, l'angiogenèse, la réponse immunitaire, l'attachement viral...

L'étude de la structure des HS, du fait de la nature flexible et hétérogène de ces molécules, a été principalement focalisée sur des analyses fragmentaires du polysaccharide au niveau des séquences d'interaction avec les protéines. Lors de ces dépolymérisations, des informations sur le polysaccharide, notamment l'épimérisation, sont perdues.

Dans ce travail, nous avons développé une approche basée sur la résonance magnétique nucléaire (RMN) bidimensionnelle ^1H - ^{13}C pour l'étude de la composition saccharidique des HS réalisée directement à partir des HS isolés de cellules marquées au ^{13}C . Pour cela, un protocole efficace de marquage et de purification des polysaccharides a été mis en place. En intégrant le volume des pics à différents déplacements chimiques par RMN, cette analyse non-destructive permet de déterminer à la fois le profil de sulfatation et d'épimérisation des HS. Cette analyse est appliquée efficacement à différents types cellulaires et est de grand intérêt pour mieux comprendre les changements dans les structures d'HS qui ont lieu lors de régulations physiologiques ou lors de développement pathologiques.

Ces résultats ont permis d'ouvrir la voie à l'analyse des HS directement au niveau des cellules par RMN du solide. Les études dans ce contexte représentent un enjeu majeur pour la compréhension des différents rôles des HS et leur capacité à interagir avec une myriade de protéines *in vivo*.

Mots-clés : Héparane Sulfate ; RMN ; Sulfatation ; Epimérisation ; Purification ; Marquage métabolique

Abstract

Glycosaminoglycans (GAGs) belong to a linear polysaccharide family which are found within all tissues, at the extracellular matrix and cell surfaces levels. Heparan Sulfates (HS) are one of the major members of this family, they are bound to a core protein to form altogether the so-called proteoglycan (PG). Depending on the localization and on the core protein, the HS – composed of a *N*-acetylglucosamine (GlcNAc) and a glucuronic acid (GlcA) [-4GlcA β 1-4GlcNAc α 1-] building block – undergo various modifications. Indeed, HS can be sulfated at different positions on both monosaccharide and the GlcA can be epimerized into an iduronic acid (IdoA). The fine structures of the polysaccharide will be able to interact with a large range of proteins and play a plethora of roles such as in inflammation processes, cell proliferation, angiogenesis, immune responses, viral attachment...

The HS structural studies, due to the flexibility and heterogeneity of the polysaccharide, have so far been restricted to HS fragments able to bind proteins. The depolymerization techniques induce valuable information losses such as epimerization.

In this work, we have successfully developed a nuclear magnetic resonance (NMR)-based approach to study HS features from ^{13}C metabolically enriched cells. For this, an effective protocol to label and purify HS has been set up. By integrating peaks' volumes at well-resolved ^1H - ^{13}C chemical shifts by NMR, the sulfation, epimerization and disaccharide profile can be determined from full-length HS. This method has been used to study HS from various cell types and is of important interest to better understand changes in HS structures that occur through physiologic and pathologic events.

The results obtained open the way to analyze HS directly at the cell surface *via* solid state NMR techniques. In this context, these studies are a major challenge to decipher the different roles of HS and their ability to interact with so many partners *in vivo*.

Keywords : Heparan Sulfate ; NMR ; Sulfation ; Epimerization ; Purification ; Metabolic-labeling